

A significant time for Europe

After languishing in recent years, the vision of a research community operating effectively at a European level is being actively pursued. Larger member states should give it their support.

European research commissioner Philippe Busquin has a dream: European scientists being free to work in labs of their choice from Sicily to Scotland, Latvia to Lisbon, being eligible for funding from any European research agency, communicating with colleagues all over the continent via high-speed research networks, and their intellectual property being protected by an affordable European patent.

This vision has long been held by the more Europhile members of the research community. Now, and not before time, the European Commission and national research ministers are beginning to rise to it, in the form of a European Research Area. Indeed, it is taking shape earlier than expected. The European science ministers' unanimous approval last week of the 'Resolution on establishing a European area of research and innovation' calls on the 15 member countries of the European Union (EU) to make it a reality without delay (see page 873).

Busquin is driven by the unarguable fact that Europe's political structure puts the continent's research community at a fundamental disadvantage compared with Japan and, especially, the United States. His concept of a cross-frontier European research space is strongly supported — in order of increasing financial clout — by the European Science Foundation, by the enthusiastic president of the European council of research ministers, Portugal's science minister José Mariano Gago, and by the heads of Europe's national research funding agencies ('EUROHORCS'). Gago and Busquin have entered the European stage at exactly the right moment; the climate for research is arguably better now than for a considerable time. Science has visibly moved up the European agenda, a fact that became palpable during the Lisbon summit in the spring. In the wake of such favourable omens and of the newly achieved power balance between the European Commission and the Council of Ministers, true reform of the European research landscape seems possible.

But Busquin knows that he can achieve little without the support

of the member states, and that his role must also be that of a persuader and catalyst. A skilful communicator, he is seizing the opportunity provided by the new open-mindedness of many government heads towards considering a joint European research policy.

But how, in concrete terms, would a young scientist benefit from the European Research Area? Most important, it would become a lot easier to gain experience in, and import skills into, any European country. At long last, mobile researchers would be relieved from a series of restrictions, such as work permits, incompatible university qualifications and complications with taxation and social security.

That will be no mean political feat. More controversy is likely to be stirred, however, by the science ministers' calls on the EU member states to progressively open their national research programmes to non-national research projects. Many believe the latter to be the crucial point with regard to true Europeanization of research. Undoubtedly, it would add a sound European dimension to many research projects, and provide them with the critical mass, if two or more national agencies were to be involved. Voluntary multilateral cooperative agreements between national agencies and research councils could be a first step in this direction, provided they agree on the necessary modification of their legal statutes and area of responsibility. It remains to be seen whether the national agencies, particularly those in the large member countries, are prepared for such a move, or whether fear of having to give away some of their cake will prevail.

Gago's enthusiasm has been prominent and based on realities. Encouragingly, at last week's ministerial council meeting, his successor Roger-Gérard Schwartzberg, the new French science minister, firmly promised to pursue the work begun by Gago and Busquin. The success of the European Research Area undoubtedly depends on his commitment, as part of the French presidency of the European Union which starts next month, and on deserved backing from Germany and the United Kingdom. ■

Welcome *Nature Immunology*

Next week sees the birth of a new Nature journal.

There are six monthly journals that carry the Nature name. The evidence suggests that all have managed to establish for themselves significant areas of top-quality publishing territory. That takes insight, enthusiasm and unrelenting effort by all involved.

The launch of the seventh Nature journal, *Nature Immunology*, has required no less. Readers will be able to judge the results for themselves on 28 June, when the first issue (dated July) goes live on the web at <http://immunol.nature.com>. As much as any other biological discipline, immunology is not only thriving but also being greatly stimulated by new developments elsewhere: genomics, informatic technologies such as microarrays, and, increasingly, computer modelling. And yet tackling the fundamental questions about the human immune system is as intellectually challenging as ever. The new journal is well set to convey this sense of excitement and to be at the forefront of such developments.

Nature's enthusiasm for publishing superb immunology will continue unabated: we wish to publish papers that are outstanding in their fundamental significance and broad interest, conveying their significance beyond the immunology community. The editorial relationship between the new journal and *Nature* itself is no different from that of the others. Publishing decisions are made by *Nature Immunology's* own staff on the basis of referees' comments: there is no consultation with *Nature* on individual papers, nor vice versa. In that sense the journals are independent. However, if *Nature* decides that an excellent paper lacks the broader significance we seek but would still be outstanding within the field of immunology, we will offer to the authors to pass the referees' comments on to *Nature Immunology* so as to save the authors effort and time.

Welcome to a significant new development for immunologists! ■





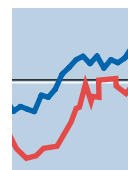
Clear vision
New plans approved
for astronomy in
Hawaii
p872



Wind of change?
UK government told
to spend more on
renewable energy
p873



Gold strike
US facility opens bid
to mimic the start of
the Universe
p874

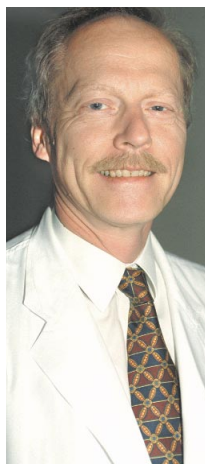


The price is right
Imminent genome
draft breathes life
into biotech shares
p875

German fraud inquiry casts a wider net of suspicion...

Alison Abbott, Munich

The fallout from Germany's most notorious case of alleged systematic scientific fraud is proving even more damaging than expected to the reputation of the country's clinical research. Attention had previously focused on a cancer researcher who had worked at the universities of Freiburg, Berlin and Ulm, but it is now suspected that other top clinical professors may have been involved in possible scientific misconduct.



Mertelsmann: data in one of his papers are under investigation.

Such suspicions were raised in a report published earlier this week by a task force that has been analysing the publications of Friedhelm Herrmann, who was first accused of fraud in 1997, and who now works as a private practitioner.

The Deutsche Forschungsgemeinschaft (DFG), Germany's research council, has followed up suspicions raised by the task force and has launched a formal investigation into a paper published in the journal *Blood* in 1994, of which Herrmann was not an author.

The task force, a small team of scientists headed by Ulf Rapp, a cell biologist from the University of Würzburg, has over the past two years been analysing the data in 347 papers published by Herrmann between 1985 and 1996. The researchers used computer programs to look for evidence of copying or manipulation of computer-stored autoradiograms, as well as photographs of suspected proteins and other molecular biological data.

It also interviewed some of the papers' co-authors in a bid to reconstruct the history of each experiment. In its report, published this week, the task force says that 94 papers include data that it thinks have been either definitely or 'highly probably' manipulated.

Fifty-three of these were published jointly

by Herrmann with his former colleague Marion Brach, who has already admitted fabricating data (see *Nature* 387, 750; 1997). Fifty-nine — including some involving Brach — were co-authored by Roland Mertelsmann, chair of the department of haematology and oncology at the University of Freiburg, in whose department Herrmann had worked.

Mertelsmann, who is well known for being the first scientist to conduct a gene-therapy trial in Germany, has accepted responsibility for the suspect publications. But he says that he was only an honorary author, and was therefore unaware of the details of experiments described in them.

Only 132 papers have been given a clean



Rapp: has analysed over 300 papers in the past two years.

bill of health by the task force. The remaining 121 were placed in a 'grey category'; a lack of cooperation from some authors, says Rapp, limited his team's access to original data, and the manipulation of data could therefore be neither proved nor disproved. Three of the five most frequent co-authors provided no information at all to the task force.

But, in an effort to gauge the extent of the problem, the task force also looked at five

...as disillusionment reigns in task force

Participation in the task force investigating allegations of scientific falsification against former cancer researcher Friedhelm Herrmann (see above) seems to have been a generally negative experience for all those involved.

Extensive evidence of data manipulation, a lack of cooperation from some authors, and the lack of interest from many of the journals in which suspicious papers had been published, disillusioned all task-force members.

The task force was set up in 1998 by the DFG, Germany's main university research funding agency (see *Nature* 395, 533; 1998). Its costs of DM750,000 (US\$370,000) were provided by the DFG and the cancer research charity the Mildred Scheel Foundation.

"At first it was very interesting," says Patrik Grünh, a postdoc on the project. "Data

falsification was a new and big thing. But later, after seeing so many falsifications, it got boring, and it makes no difference to the principle involved to prove that 94 papers, rather than 84, are suspect."

For the leader of the task force, Ulf Rapp of the University of Würzburg, the most frustrating thing was the "moving target" provided by the scientists under investigation. "Everyone pointed their finger at someone else when asked who was responsible for each experiment," he says.

Even given the lack of cooperation from some authors, the extensive 'grey category' of papers, where no strong decision on whether the data quoted might have been tampered with, was also a sign that clinical laboratories require mechanisms for ensuring good scientific record-keeping, he says.

Rapp blames lax standards in clinical research in Germany for

the outbreaks of scientific misconduct. He attributes this largely to the fact that clinicians do not have to be formally trained in research methods. This problem is exacerbated by a requirement for clinicians to have a long publication record in order to achieve promotion.

Rapp is also angry that it has so far proved impossible for the courts to bring a legal case against either Herrmann or Marion Brach, and he believes that professional societies have not been sufficiently vociferous in condemning misconduct (see *Nature* 395, 532–533; 1998).

But in response, Volker Diehl, president of the German Society for Haematology and Oncology, says the society responded rapidly, expelling both Herrmann and Brach, and setting up a group of retired professors to whom young scientists can turn if they are concerned about misconduct in their laboratory.

A. A.

randomly selected papers on which Mertelsmann — but not Herrmann — was listed as a co-author. It asked the authors for original patient data to allow them to analyse tables and graphs.

According to Rapp, sufficient original data were provided to fully analyse only one paper, published in *Blood* in 1994 (volume 84, 1421). This paper investigated a technique to help cancer patients recover from the bone-marrow damage that reduces blood-cell counts following high-dose chemotherapy.

The paper looked at the rate of recovery of blood cells when the patients' own peripheral-blood progenitor cells, taken before chemotherapy, were transplanted back, either unseparated or purified. The paper concluded that recovery was rapid and complete in either case.

But the task force's report says the paper contained many "irregularities and indications that data had been improperly handled". Rapp told *Nature* that the published graphs describing blood-cell recovery omitted data on a significant proportion of the patients. "If you mis-portray clinical data in important journals you may encourage others to adopt a practice that could put patients at risk," he says.

In response to the task force's findings, the DFG has launched a formal investigation of Mertelsmann and of Lothar Kanz and Wolfram Brugger, respectively professor and senior researcher at the University of Tübingen, and co-authors of the *Blood* paper.

The University of Freiburg has also opened an internal investigation into the allegations. Kanz declined this week through his lawyers to comment on the investigation; the two other authors did not respond to several requests from *Nature* to discuss the issues raised by the task force in its report. ■



Looking up: expansion plans for the facilities at Mauna Kea have been welcomed by astronomers.

Hawaii reaches for the stars with astronomy master plan

Rex Dalton & Alison Abbott

A long-awaited master plan for the future development of astronomy projects on Mauna Kea in Hawaii was approved last week by the University of Hawaii. The agreement comes after two years of negotiations with native Hawaiian cultural groups and environmentalists and a last-minute decision to drop two of the planned facilities.

At the same time, the university has announced that its three-year search for a new director for its Institute of Astronomy is over. Rolf-Peter Kudritzki, currently director of the Institute for Astronomy and Astrophysics at the University of Munich, will take up his new post in October.

Kudritzki says that Mauna Kea has "enormous scientific potential because it is host to so many international observatories".

The master plan to set up an 'astronomy precinct' within the Mauna Kea Science Reserve, and the recruitment of Kudritzki, will allow Hawaii to move forward with a number of plans for expansion and improvements at Mauna Kea. The site is currently home to a number of major

facilities, such as the Keck telescope.

Under the plan, three new facilities will be built, five of the existing 13 facilities will be redeveloped, and two of the existing facilities will be expanded.

Native Hawaiians have expressed concern that telescopes and other facilities built near Mauna Kea's 14,000-foot-tall peak would disturb land on which traditional gods are believed to live. Environmentalists have raised concerns about possible damage.

As part of the compromise agreement worked out, the university has cancelled plans for a 4- to 12-metre telescope and an optical/infrared interferometer facility.

In response to local concerns, the master plan calls for an advisory group of native Hawaiians to have a say in future expansion. But native Hawaiians themselves remain sharply divided on Mauna Kea astronomy development. Some leaders have endorsed the master plan, but others continue to threaten legal action to block it.

Kudritzki says the "master plan presents an appropriate balance between the concerns of ecology, religion and astronomy". ■

Oxford epidemiologist wins apology for promotion slur

Natasha Loder, London

Sunetra Gupta, the University of Oxford epidemiologist wrongly accused of having won support for a readership post through a relationship with her head of department (see *Nature* 403, 353; 2000), has received a retraction and an apology from Roy Anderson, the chair of the appointment committee.

Anderson wrote to Gupta last week unreservedly withdrawing an allegation for which he said there was "no foundation in truth whatsoever and which I accept I never should have made". It has taken Gupta eight months to get the retraction, during which

time, she says, she received little support from the university, which she felt had tried to brush the situation under the carpet.



Gupta: waited eight months for apology.

Gupta says that Anderson's comments about her were "totally unacceptable and could not be allowed". She rejects as "highly offensive" criticism from some members of the department of zoology who say she was manipulated into complaining.

The university has already upheld complaints that Anderson had intimidated members of Gupta's appointment committee. His behaviour also led to complaints about his management of the Wellcome Trust Centre for the Epidemiology of Infectious Diseases (see *Nature* 403, 695; 2000), and ultimately his resignation from the centre, from Oxford University and as governor of the Wellcome Trust.

In his letter to Gupta, Anderson says he was under "a great deal of personal stress at the time". Anderson has also agreed to pay Gupta's legal costs and to make a donation to charity. ■

Europe edges closer to an integrated science policy

Quirin Schiermeier, Munich

The concept of a single 'European research area' last week took another step towards becoming a reality. At a meeting in Luxembourg, the research ministers of the 15 member states of the European Union (EU) fixed a schedule for changes intended to increase the efficiency and competitiveness of basic science in Europe.

The aim is to forge closer links, at both practical and strategic levels, between the scientific activities and science policies pursued by individual member states through their national research agencies.

A number of scientific organizations welcomed the agreement as a major achievement. "What we have here is a clear outline of a true, and urgently needed, joint European science policy," says Enric Banda, secretary general of the European Science Foundation (ESF).

The various measures for establishing the single research area are now set to be implemented by the end of 2001. They include creating a single European patent, compatible benchmarking of national research systems, building a high-speed trans-European research network, removing bureaucratic obstacles to the mobility of researchers, and opening national research programmes to selected transnational research projects.

The resolution is based on a proposal put forward in January by Philippe Busquin, the European Commissioner for research. However, as expected, a second resolution calling on the commission to support the infrastructure costs of Europe-wide life-science facilities was not endorsed (see *Nature* 405, 723; 2000).

Now that the European research ministers have formally endorsed Busquin's draft, the integration of national and EU research programmes appears to have become a significant practical priority.

Reinder van Duinen, president of the ESF and head of the Netherlands' national research council, says that science and technology are now high on the agenda of Europe's leaders. He is particularly pleased that the research ministers have agreed on well-defined tasks and a strict timetable for their realization. "The fact that there is now a managerial schedule puts European science policy on the same political level as the monetary union, defence or foreign policy," he says.

Richard Brook, chief executive of the British Engineering and Physical Sciences Research Council, similarly applauds Busquin's role in promoting the convergence of Europe's national science bases. He sees the resolution's suggestion of a coherent and comparable system of benchmarking nation-



al research performances as a necessary tool for achieving the common research space.

But Brook is also keen to stress that regional research efforts and small-scale collaboration will remain important. Moreover, he points out that, on legal and administrative grounds, national research councils have only limited opportunities to make funds available Europe-wide.

Even if such chances increase, there will be no dramatic overnight change. Indeed, the resolution suggests that the networking of national research programmes should start on a voluntary basis for goals with a European dimension. According to Banda, this matches the objectives of the ESF, which has recently set up its EUROCORES programme to aid collaboration between agencies or research organizations from different countries.

Ernst-Ludwig Winnacker, president of the DFG, Germany's main funding organization for university research, calls the European research area an extremely helpful concept. "Increased networking between national funding agencies is really desirable," he says. "The DFG will do everything it can to help make Busquin's effort a success."

At a press conference following last week's research council meeting, Busquin emphasized the advances in science and technology policy that have been made during the past six months. "Science has finally taken the place in Europe it deserves," he said.

France's new science minister, Roger-Gérard Schwartzberg, will next month succeed José Mariano Gago, the Portuguese science minister, as president of the European research council — the body that brings together EU research ministers. Schwartzberg said last week that he is "fully sympathetic" with the course mapped out by Gago and Busquin.

► <http://europa.eu.int/comm/research/area.html>

► <http://www.esf.org/about/eurocores.htm>

UK panel calls for more cuts to carbon dioxide emission

Natasha Loder, London

Britain should reduce its emissions of carbon dioxide by 60% over the next 50 years, and set an international model for tackling climate change, according to a key independent report to the UK government published last week.

The Royal Commission on Environmental Pollution, in a report entitled *Energy — The Changing Climate*, says the United Kingdom has a "moral imperative" to act now to curb emissions. But it also doubts whether the government will be able to meet its target of a 20% reduction in carbon dioxide emissions in the next decade.

The commission is an independent body that reports to parliament on the crucial environmental issues facing Britain and the world. Its report also reveals that the United Kingdom spends less on energy-related research than almost any other developed nation, and warns the government that it must reverse its trend of cutting spending on such research if new energy systems are to be developed.

"We cannot expect other nations to do their part in countering this threat — least of all if they're much less wealthy — unless we demonstrate we are really serious," says Tom Blundell, head of the department of biochemistry at the University of Cambridge and the commission's chairman. According to Blundell, the UK environment minister, Michael Meacher, agrees that a 60% reduction is needed in the long term.

Blundell says that more research should be focused on the electricity grid to ensure the effective distribution of high-quality energy, including that of intermittent energy sources such as solar and wind power. The report says Britain lags far behind the rest of Europe in developing renewable-energy technologies.

The report coincides with a new round of discussions in Bonn over the detailed implementation of the Kyoto Protocol on reducing the emission of greenhouse gases. Effective mechanisms for reducing emissions must be finalized in time for a crunch meeting in The Hague this November.

Last week, three environmental groups claimed that some industrialized countries were trying to negotiate looser terms for certain mechanisms in the treaty that would

undermine its effectiveness.

A number of commentators point out the irony that many of the hotly debated flexible mechanisms are there at the insistence of the United States, which is highly unlikely to ratify the Kyoto Protocol.

The November meeting coincides with the US elections, and the US position is not expected to have changed by then. However, the United States does not have a blocking vote, so the protocol can be implemented without it. The new administration could still instigate mitigation measures if it chose to do so.

But observers point out that Kyoto's 5.2% target would not be reached without US involvement. Benito Muller, a senior research fellow from the Oxford Institute for Energy Studies, says that although this is "neither here nor there", more substantial steps are needed. "Industrialized nations have committed themselves to showing leadership [on global warming] and Kyoto is the tool by which we can do it," he says.

A continuing source of controversy involves a mechanism that would allow industrialized nations to claim credits against their emissions by investing in clean technologies — such as nuclear power — in developing countries.

The nuclear industry is emphasizing the major role that nuclear energy could play in reducing global warming. But critics, including Britain's Royal Commission, warn that the issue of nuclear waste must be dealt with before it can qualify as a clean technology. ■

► <http://www.rcep.org.uk/newenergy.html>



Clean power: wind turbines, such as these in Jutland, Denmark, are a rare sight in Britain.

Brookhaven collider opens its quest for Big Bang conditions

Steve Nadis, Boston

Two gold nuclei, hurtling in opposite directions around a 2.4-mile ring, crashed into each other last week, unleashing 10×10^{12} electron volts of energy. This startling event marked the debut of the Relativistic Heavy Ion Collider (RHIC), as it came online at Brookhaven National Laboratory in New York.

Over the next few weeks, scientists hope to increase the accelerator's power to generate collisions releasing about 40×10^{12} electron volts of energy. The aim is to simulate the conditions that existed during the first few microseconds after the Big Bang.

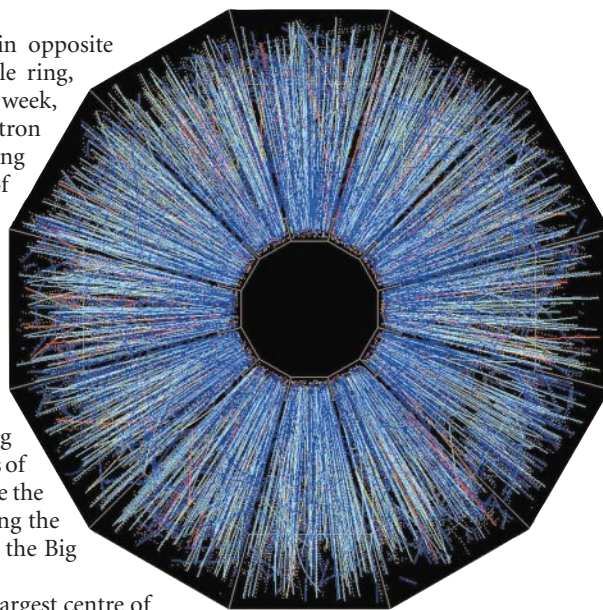
"RHIC can produce the largest centre of mass energy achieved by any accelerator, which makes it the best machine for studying the birth of nuclear matter in the Universe," says Tom Kirk, Brookhaven's associate director for high-energy and nuclear physics.

Researchers at the collider are trying to create a plasma of unattached quarks and gluons — the state of matter thought to have existed just after the Big Bang. They want to study the 'phase transition' that occurs as these particles combine to form protons and neutrons.

In February, scientists using the Super Proton Synchrotron (SPS) at the European Laboratory for Particle Physics (CERN) in Geneva announced they had found "compelling evidence" for the quark–gluon plasma (see *Nature* 403, 581; 2000). But the SPS did not generate enough energy for researchers to see the plasma directly.

The RHIC, which is about 10 times more powerful, "should produce the energy unequivocally needed to see the plasma", says Kirk. The collider will yield temperatures well above one trillion degrees kelvin at the site of collision.

RHIC scientists will scour the rubble left by the ion collisions for clues to the plasma's existence. One tell-tale sign would be gamma-ray radiation produced by interactions of unconfined, high-energy quarks in the plasma. "That would certainly be a smoking gun," says Brookhaven physicist Tom Ludlam. Other direct evidence of the plasma's existence could come from electrons emerging from the plasma before it cools and reverts to ordinary nuclear matter. This cooling is very quick —



Collider scope? An end view of a collision of two gold nuclei at Brookhaven's RHIC collider.

the plasma is expected to last about 10^{-23} seconds.

The researchers also hope to find out the precise value of the critical temperature at which the quarks in protons and neutrons become unconfined — or, conversely, the temperature at which free quarks coalesce into nucleons. And by measuring specific properties of the quark–gluon plasma, such as its temperature, density and entropy, they plan to test key predictions of quantum chromodynamics, the theory describing the strong nuclear force that joins quarks with gluons.

The RHIC machine was switched on after a concerted effort by Brookhaven officials to allay public fears of doomsday scenarios stirred up by alarmist news coverage. These media accounts suggested that the dense plasma created in the RHIC's high-energy collisions could produce a miniature black hole or a previously unseen form of nuclear matter called 'strange matter' — either of which might, according to the scare stories, quickly devour the planet. A panel of physicists dismissed these concerns in a report issued last autumn.

The current experiment at the RHIC is scheduled to run until August. The machine will eventually be superseded by the 30-times more powerful A Large Ion Collider Experiment (ALICE) at CERN's Large Hadron Collider, which is due to be started up in 2005. "But there is a lot to learn before ALICE is turned on," says Ludlam. "This reality is unexplored territory." ■

Shares rebound at scent of draft sequence

Paul Smaglik, Washington

The world's scientific community may still have to wait a few more days for the news — widely expected early next week — that the draft sequence of the human genome has been completed. But shares in genomics companies have already been rising steadily for several weeks in anticipation.

These shares have been through a roller-coaster ride in the past six months. They began to climb steadily last December, immediately before the announcement that human chromosome 22 had been sequenced. After that milestone, prices were propelled further upwards by word that both the publicly funded Human Genome Project (HGP) and its commercial counterpart Celera Genomics of Rockville, Maryland, were speeding through the human gene sequence (see *Nature* 403, 119; 2000).

But the bubble burst in early March, and the decline gathered pace when US President Bill Clinton and British Prime Minister Tony Blair issued a joint statement that genomic information should be freely available. This was widely interpreted by investors as an attack on companies that intended to patent genes, and contributed directly to the second-biggest drop in the history of NASDAQ (see *Nature* 404, 424; 2000).

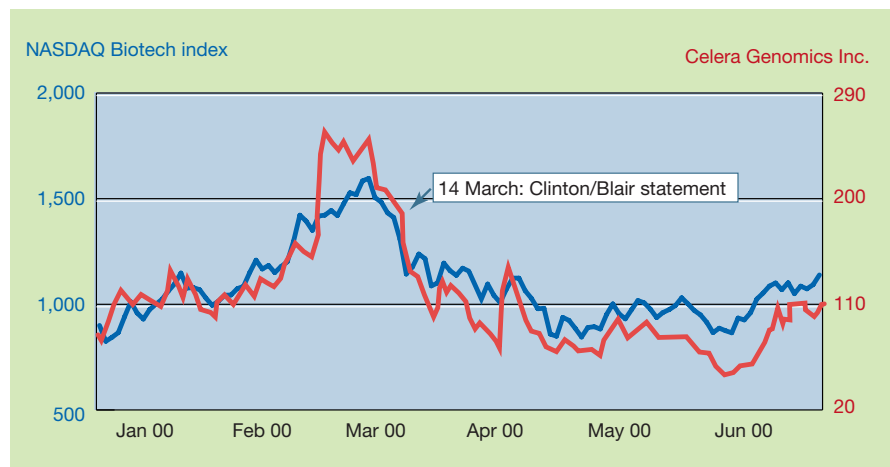
It is likely to be a long time before the peak values are regained. But anticipation of the impending sequencing announcement seems to have repaired some of the damage. For example, the value of a Celera share plummeted from \$276 at the beginning of March to \$122.5 immediately after the Clinton/Blair announcement.

Celera's shares continued to slide, hovering around the \$70 mark until early June. Now that both Celera and the HGP are due to announce that they have drafts of the human genome in hand, the price has climbed back above \$100 a share (see figure).

Shares in Incyte Genomics of Palo Alto, California, were trading on Monday at around \$92, after dipping to under \$50 following the announcement. The restoration in the fortunes of Human Genome Sciences of Rockville, Maryland, has been even more dramatic; its shares, which fell from a peak of \$232 to just above \$50 a share following the Clinton/Blair statement, are now back to around \$135 a share.

"The stocks are probably still less than half of the highs of earlier in the year, but they've certainly rebounded from the lows," says Winton Gibbons, an analyst from William Blair & Co. in Chicago.

Gibbons attributes some of the rebound to anticipation over the next stage of the genome project, but says that other factors may also be at work. For example, he says, investors soon realized that the Clinton/Blair



Bouncing back: share prices are on their way up again.

statement was relatively toothless. Guidelines issued by the US Patent and Trademark Office showed that the government wasn't going to restrict gene patenting as much as some people had believed.

Gibbons expects the rebound to benefit 'next-generation' companies as well. For example, news of plans being made to sequence the genomes of other organisms, such as mice, will help functional genomics companies, which analyse animal genes that have human homologues.

The March panic sell-off was a blessing for bargain hunters, says David Molowa, an analyst with Chase H&Q in New York. The aftermath resulted in a 'fire sale', Molowa believes. "Folks with a little bit of guts

stormed in and started reversing the trend."

Both Molowa and Eric Schmidt, an analyst with SG Cowen Securities in New York, predict that there will now be a period of consolidation, with large companies buying smaller, specialized ones. "There are too many genomics companies with too narrow a piece of the pie," says Schmidt.

If genomics companies are to maintain or increase their value, they will have to find a way to make money beyond the stock market. All have different approaches. The HGP is becoming a conventional biotech company to develop drugs, while Celera is sticking to its database subscription model. Until companies turn a profit, it will be hard to assess their true value, Schmidt concludes. ■

GM food 'dumped on India as aid'

K. S. Jayaraman, New Delhi

A Delhi-based environmentalist group has accused the United States of 'dumping' genetically modified (GM) soya and corn on India in the form of food aid "in a bid to create an artificial market and keep the tottering genetic-engineering giants afloat".

The United States began sending food to Orissa in November 1999. The US Agency for International Development does not deny that it supplies GM food in aid packages, saying that it does not distinguish between GM and non-GM food.

Vandana Shiva, president of the Research Foundation for Science, Technology and Ecology (RFSTE), says that the \$4.2 million US food aid sent to cyclone victims in the state of Orissa last October contained genetically modified forms of soy beans and corn rejected by traditional markets in Europe and Japan.

The discovery was revealed by test

reports from Genetic ID, an American company that tests for GM organisms in food. The company tested samples of food that had been collected from Orissa and sent for analysis by RFSTE, and concluded that "significant levels of genetically modified DNA were detected in the samples".

India has no policy on GM foods at present. The secretary of the Department of Biotechnology, Manju Sharma, says that the department is still "in the process of finalizing" the policy in consultation with the ministries of health, the environment, food and agriculture.

Shiva complains that, in the absence of policy and test facilities, India has become an illegal dumping ground for unwanted GM foods. She claims that the US government "used the money intended for relief to the poor to subsidize the biotech industry, in order to create market entry for GM products". ■

US science R&D funding favours fifteen states

Washington Fifteen US states receive 80% of federal R&D dollars, with California, which received \$14.4 billion in 1998, at the head of the pack, according to a report released last week by the Rand Corporation. The report, which claims to be the first to identify and describe virtually all federally funded R&D activities within the states, was prepared for the White House Office of Science and Technology Policy.

Overall, says the report, more than \$80 billion a year is being spent on scientific R&D by the US government. In Maryland, 34% of the total federal funding came in support of R&D. But despite the concentration of funds, the report points out that for many smaller states, such as Rhode Island and New Hampshire, federal R&D funding on a per-capita basis is higher than in larger states receiving more federal R&D funds, such as Texas and Florida.

► <http://www.rand.org/hot/Press/FederalRD.html>

Russian ecologists in arms over vanishing ministries

Moscow Russian ecologists from more than 1,500 different professional organizations have united for the first time to protest against the disappearance of two ministries — ecology and forestry — from the new cabinet. They have also complained about the recent disbandment of the 'ecological police' by the ministry of home affairs, and the elimination of compulsory ecology courses in schools.

Boris Yatskevich, the minister of natural resources, said at a national conference on conservation last week that the cabinet's reorganization will help industry, which is suffering from "endless ecological regulations". But the 368 delegates at the conference sent a message to President Vladimir Putin and the Russian parliament arguing that the newly formed Ministry of Natural Resources should not be allowed both to exploit nature and to regulate this process.

UK research council comes under review

London The British government is to carry out a review of the activities of the Council for the Central Laboratory of the Research Councils (CCLRC), which provides large research facilities for UK science such as the Rutherford Appleton Laboratory near Oxford. The government has said that all relevant options will be considered, including abolition, privatization or "strategic contracting out".

The review will look at the charter, mission, contributions and past performance of the CCLRC. If the first stage

of review confirms the council's continued operation, a second stage will look at how its performance could be improved. The council, which employs 1,700 staff and has an annual turnover of £100 million (US\$150 million), was set up in 1995 when the former Science and Engineering Research Council was split up.

University open to offers for work on nuclear reactor

San Diego The University of California at Davis is seeking academic and commercial research proposals for a nuclear reactor recently acquired from the US Air Force (USAF). Materials science, biomedical research, isotope studies, nuclear engineering and neutron activation analysis in archaeology, anthropology, geochemistry and immunology are potential areas of study.

The university has \$500,000 in research funds available for studies at the neutron beam reactor, control of which was transferred earlier this year from the USAF (see *Nature* 316, 401; 1999). The reactor, located at the USAF's McClellan Air Force Base, about 20 miles from the university, was built in 1990.

Police extend scope of nerve-gas test inquiry

London Following hundreds of complaints about deaths and illnesses, British police are widening their investigations into chemical-warfare tests allegedly carried out on humans during the 1950s at a Ministry of Defence research centre. The inquiry into the Porton Down research centre is being led by local police and detectives from the ministry.

The investigation began when former servicemen alleged that they were tricked into volunteering for the tests, which subsequently damaged their health. The inquiry now involves at least one allegation of murder. The Wiltshire police are appealing to the government for help in funding the rising costs of the inquiry, which they say could last until the end of the year.

Radio astronomers increase their allotted frequencies

Istanbul After three years of planning by radio astronomers, delegates to the World Radio-communication Conference, which has just ended a month-long meeting in Istanbul, Turkey, have approved the first new dedicated frequency allocations for radio astronomy since 1979, when millimetre-wave astronomy was in its infancy.

The conference agreed to provide protection for all the frequencies between 71 and 275 gigahertz (GHz) currently used by radio astronomers, adding more than 90 GHz of spectrum to the 44 GHz already set



Radio telescopes are boosted by frequency gain.

aside in this frequency range. "It's a win for millimetre-wave science," says John Whiteoak of the Australia Telescope National Facility. "This secures its future."

Los Alamos breathes a sigh of relief

San Diego Two computer hard drives loaded with international nuclear-weapon secrets that had been embarrassingly 'lost' by the Los Alamos National Laboratory (see *Nature* 405, 725; 2000) mysteriously reappeared last week when they were found behind a photocopier in a secure area.

The circumstances surrounding the discovery of the missing hard drives in an area previously searched prompted some authorities to suspect that one of the scientists with access to the X Division area surreptitiously returned them. The events prompted some Congressmen to call for the resignation of the Department of Energy secretary, Bill Richardson. He rejected the call, saying he is trying to balance the needs of science and security at the laboratory.

Photon project harnesses high-schools' PC power

Munich Pupils at 73 high schools in the German state of Baden-Württemberg, as well as 1,100 individuals, are making their personal computers (PCs) available online to a simulation of the spreading of X-ray photons set up by the University of Tübingen. The project has been launched by the state science ministry to increase the public understanding of astrophysical research, but will also assist astrophysicists in carrying out simulations aimed at understanding the nature of X-ray pulsars.

The project, called XPulsar@home, is modelled on the successful SETI project in the United States, which makes use of the unused computer capacity of around 950,000 PCs to search for extraterrestrial intelligence (see *Nature* 400, 804; 1999). Pupils' participation will be kept alive by a competition between high schools for the highest number of simulations carried out.

A question of trust

It is the world's biggest medical research charity and it exerts a huge influence over UK science policy. But is the Wellcome Trust becoming a victim of its own success, asks Natasha Loder.

The Wellcome Trust can rightly claim to be the saviour of British biomedicine. Without its largesse, Britain would almost certainly have ended the 1990s as a second-rate biomedical power. Instead, labs benefiting from Wellcome funding have prospered scientifically, and the trust's emphasis on improving salaries and career structure for researchers has shamed the UK government into action. What is more, by pumping millions of pounds into the Sanger Centre, near Cambridge, the Wellcome Trust has become a significant force in the international Human Genome Project. The trust's spending now approaches double that of the government-funded Medical Research Council.

But over the past year, the shine has started to come off the trust's image. When its plans to extend the genome campus that it has built around the Sanger Centre fell foul of restrictions on development in rural areas, the trust was accused of arrogance in its dealings with local people (see *Nature* **400**, 803; 1999). Then its partnership with the government in a £200 million (US\$302.5 million) project to build a powerful new synchrotron radiation source ended in an unseemly row. The decision to build the facility at the trust's preferred site of the Rutherford Appleton Laboratory, near Oxford, rather than at the site of Britain's existing synchrotron at Daresbury, near Manchester, ignited a political controversy. Ministers sought to pin responsibility for the decision onto the trust, which rejected that version of events.

But perhaps most damaging was the downfall of Roy Anderson, a leading epidemiologist, and one of the governors who run the trust. Anderson, who also directed the Wellcome Trust Centre for the Epidemiology of Infectious Diseases at Oxford University, was forced to resign from the university, the centre and as a trust gover-



Management problems at Wellcome's Oxford epidemiology centre damaged the trust's public image.

nor, following the publication of two damning reports into his management of the Oxford centre. Initially triggered by allegations that Anderson had made sexual slurs against a female colleague chosen for a senior position at Oxford, the investigations went on to examine his failure fully to disclose business interests that had become entangled with the Wellcome centre's research activities (see *Nature* **404**, 696 & 802; 2000).

In the wake of these bruising events, some of the trust's friends are urging it to consider internal reforms to safeguard its reputation. In some ways, the Wellcome Trust is falling victim to its own success. Through shrewd

management of its investments, the trust's spending power and influence have grown phenomenally over the past eight years (see 'The rise and rise of Wellcome', opposite). But governing structures and procedures that were adequate for a modest-sized charity may not be appropriate for a behemoth that — by working with government as an equal partner — helps direct national science policy.

David Lane, a leading cancer geneticist at the University of Dundee, offers a single comment: "The Wellcome Trust has had a huge beneficial impact. Going forward, the key concern will be governance because they are an independent body but have the responsibility for a large fraction of UK research."

Some senior scientists familiar with the trust's workings think it needs to adopt a more professional approach to public and government relations. But most important, now the Anderson affair has thrown a spotlight onto the role of the trust's governors, is addressing concerns about transparency and accountability at the top of the organization.

In unguarded moments, some British biomedical scientists complain that the trust's governors have too much power, and voice



Success: the Sanger Centre is a major player in the Human Genome Project, thanks to the trust's millions.



Ogilvie: the trust needs to be more transparent.

director from 1991 until 1998. "But as the trust has grown, it has to deal with the perception that governors may have conflicts of interest. It's about transparency."

Scientists are unwilling to put criticisms on the record. Wellcome Trust funding is so important in Britain that nobody is prepared to bite the hand that feeds them. "This is anonymous. I don't want to get my funding cut off," one biomedical scientist told *Nature*. Says another: "Why am I going off the record? That is part of the problem." Although scientists are unanimous in praising the trust's contribution to British biomedical research, some are troubled that such a powerful body does not seem to be accountable in the same way as government research agencies. These have to report to politicians, who in turn must be elected. With the Wellcome Trust, control resides firmly in the hands of its governors.

A position of power

These governors are the charity's trustees, and they manage the trust set up by the will of pharmaceuticals mogul Henry Wellcome. This endowed the organization to support scientific research conducive "to the improvement of the physical conditions of mankind". Although this leaves the trust free to support a variety of disciplines, it has focused on biomedicine. Unusually for the trustees of a British charity, the governors receive a salary, currently between £50,000 and £80,000 a year. Wellcome's will specified that trustees should be paid, and over the years the governors have gone to court to update the value of this remuneration. Of the seven current governors, all but one — the chair, who is an industrialist — are leading biomedical scientists. Two work in Oxford, two in London and two in Cambridge (see 'Who's who', overleaf).

Although the day-to-day running of the Wellcome Trust is left to its director, Mike Dexter, he says that each governor devotes about two-and-a-half days a week to the organization. They make investment decisions, determine overall policy, sit on committees that consider grant applications, take yearly decisions about how much money will be

available for each subject area, and appoint new governors. Their laboratories are also eligible for Wellcome Trust funding — and in some cases, governors head institutes that are financed, in large part, by the trust.

It is this combination that raises concerns about conflicts of interest. The governors could, for instance, decide to spend a large proportion of trust money on a subject area in which one or more of them has a research interest, or even directs an institute. Procedures exist to deal with such conflicts, says Dexter. Governors withdraw from committees when their own grants are being discussed. Likewise, when Oxford University matters come up during a meeting,



Dexter: preparing a new constitution.

the Oxford-based governors have to retire.

Diana Garnham, general secretary of the Association of Medical Research Charities — an umbrella organization to which the trust belongs — says that her group monitors the amount of money given to the governors and has never found "anything in particular to worry about". Although a regional

breakdown of the trust's funding reveals an apparent bias towards London (see Figure, overleaf), this can be explained by the capital's strength in biomedical research.

But given the enormous sums of money being distributed by the trust, some researchers feel that further action is needed to address perceptions of conflicts of interest. The scientific calibre of the Wellcome governors is such that it is not surprising that they and their associates should attract trust funding, says one senior biomedical scientist. "But this is a potential problem and is very difficult to deal with." And the Anderson affair did raise a warning: if a governor's behaviour does depart from the highest standards, it may be difficult to address the issue before matters have got out of hand.

Time for a make over

The Wellcome Trust seems to be concerned about the dents its image has taken, and has hired a public relations company to find out what people think about it. Trish Evans, currently communications director at the National Society for the Prevention of Cruelty to Children, will take over as head of communications for the trust in July. Part of her brief will be to improve its responsiveness to the media. She is also an expert on relations with government.

The rise and rise of Wellcome

For more than four decades, the Wellcome Trust was a solid, unspectacular performer. When Henry Wellcome died in 1936, his will ordered that the shares of the pharmaceutical company he founded — the Wellcome Foundation — be put into trust to support non-commercial research. The company continued to make a profit, which the trust ploughed into research at academic labs, particularly in disciplines that attracted scant government funds.

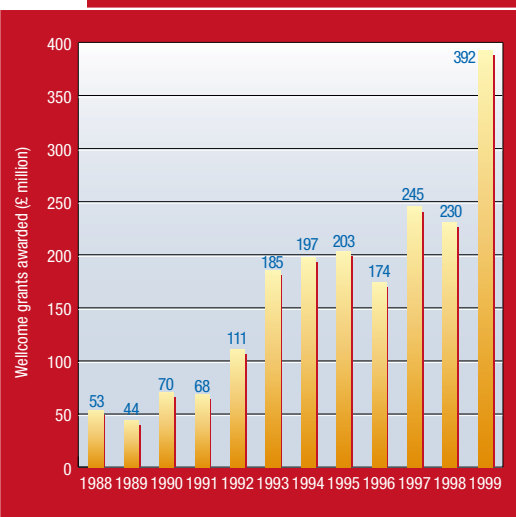
In 1986, with the Wellcome Trust's annual spending standing at around £20 million, its

trustees made a move that sent the trust down the road to greatness. Realizing that the trust could earn more by spreading its investments, they floated the drugs company on the stock exchange and sold a fifth of its shares. That year the trust awarded the first of its five-year programme grants — marking a move into support for long-term research projects.

But the most profound change came in 1992, when the charity sold a further tranche of shares as the Wellcome company merged with Glaxo. This raised £2.3 billion, and the trust's research budget more than doubled overnight. Since then, thanks to a buoyant stock market and sound financial management, the trust's asset base and spending have continued to grow (see Figure).

Improving career structure and pay for scientists was always a priority for the trust. It has led the way in boosting stipends for PhD students, and in the past year it has bumped up the pay of 270 of its young scientists by 30 per cent, challenging the UK government to do the same.

As its wealth has grown, the Wellcome Trust has moved into larger projects, in some cases providing core support for research institutes. Most recently, the trust has entered into an unprecedented partnership with the government. In July 1998, it agreed to contribute £300 million to a £750 million Joint Infrastructure Fund, which will pay for large items of research equipment for British academics.



But some observers feel that more fundamental changes are needed, and Ogilvie has some specific ideas. To increase transparency, she suggests, the trust's annual report and website should contain clear information on how governors are selected, the terms of their employment, and how they deal with conflicts of interest. In appointing new governors, she adds, there is a case for recruiting scientists from a wider variety of backgrounds, to dispel the 'old boy' perception. "They should make an effort to be more representative of the scientific community," says Ogilvie.

Shortly before *Nature* went to press, the trust announced the appointment of three new governors, who will serve from October. This will widen the geographic distribution of the governors, as one is based in Edinburgh. The new appointments also add a solitary woman to the board of governors.

But viewed from the United States, this does not go far enough. The trust should rethink the policy allowing its governors to receive substantial funding from the organization, says Maxwell Cowan. Until March this year, Cowan was vice-president of the Howard Hughes Medical Institute (HHMI) — the nearest US equivalent to the Wellcome Trust. "We don't have a situation where an individual in receipt of a grant is also determining how funding should be distributed," he adds.

Charities and foundations in the United States go to great lengths to ensure that trustees cannot benefit from money that the



Cadbury: content with the status quo.

bodies distribute. "It is called self-dealing and everyone here is very hyper about it," says Bob Potter, the HHMI's director of communications. The HHMI's board is largely made up of business executives, and its two scientist members do not get involved in individual funding decisions — as board members they merely vote on whether to accept or reject the final list of awards recommended by the HHMI's scientific committees.

Differences between the Wellcome Trust and the HHMI, and between Britain and the United States, make it difficult to draw direct parallels. The bulk of the HHMI's awards support elite 'Hughes investigators', freeing them from the treadmill of grant applications. The Wellcome Trust, meanwhile, funds a wider range of activities. And although it spends similar amounts to the Wellcome Trust, the HHMI's budget is dwarfed by that of the US National Institutes of Health. This, and the huge size of the US biomedical research community, makes it easier to avoid placing individuals in positions of conflicts of interest. The concentration of British scientific talent into a small number of leading universities, says Potter, "seems to call for creative thinking about how to avoid self-dealing — and even the appearance of self-dealing".

Transparent thoughts

Ultimately, that is the responsibility of Dominic Cadbury, chief executive and chairman of the food and drinks manufacturer Cadbury Schweppes, who has chaired the Wellcome Trust's board of governors for the past six months. Unfortunately for those who favour reform, he appears to be content with the status quo. "I haven't come with any brief to change things," says Cadbury. "The force for change should come from the director."

Dexter has been working on a new constitution and various other policy documents, drafts of which sit prominently on Cadbury's desk. But the trust's director is hired by, and reports to, the governors. And Cadbury appears less open than Dexter to the idea of change. One suggestion to improve transparency, made by Frank Karel, vice-president for communications at the Robert Wood Johnson Foundation, a leading US philanthropic organization in health and healthcare, is that minutes of governors'

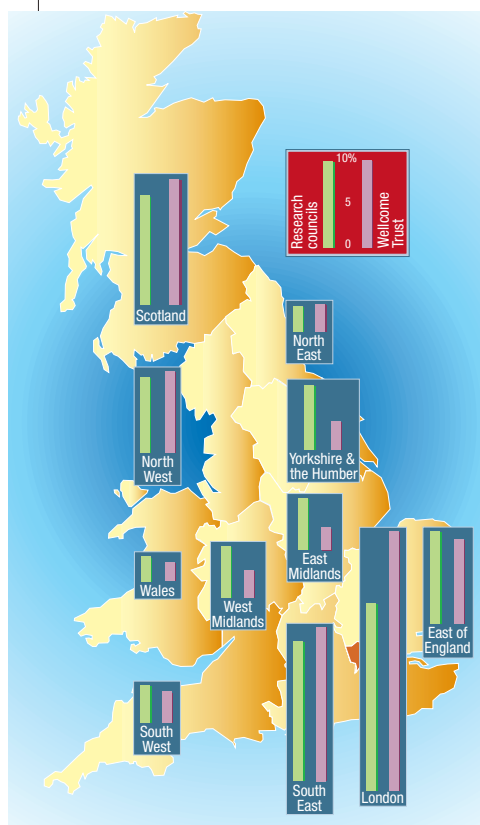
meetings should routinely be made public. Dexter describes the idea as "interesting". But Cadbury says: "I don't think so", adding that this would never happen in a company.

Cadbury is happy with the trust's existing provisions for managing potential conflicts of interest among the governors. "The two principles are: transparency when there is a conflict and that the person should be no part of the decision," he says. But this relies on the governors declaring their interests to the trust, which is rather less transparent to those outside the organization.

Cadbury also rejects the idea of barring governors from receiving research funding from the trust. "That sounds to me like you are denying that their work is important," he says. "It seems to me absurd." The Wellcome Trust is such an important source of funds in British biomedicine, he argues, that adopting the US approach would be too restrictive.

But that is the crux of the problem. Having grown so quickly to become such a big fish in a relatively small pond, avoiding the appearance of conflict of interest is a difficult task. "It is challenging to think what the right way forward is," says one scientist. The trust's friends hope that Cadbury and his colleagues are up to that challenge, so that future headlines are not about the trust's embarrassments, but its scientific achievements. ■

Natasha Loder reports from London for *Nature*.



Regional accents: percentage breakdown of total spend by UK research councils and the trust.

Who's who

In addition to the chair, industrialist **Dominic Cadbury**, the Wellcome Trust's governors are:

Michael Rutter, professor of psychopathology at the Institute of Psychiatry in London

Julian Jack, professor of physiology at the University of Oxford

David Weatherall, honorary director of the MRC Molecular Haematology Unit and honorary director of the Institute of Molecular Medicine, both at Oxford University

Christopher Edwards, an endocrinologist and principal of Imperial College School of Medicine, London

John Gurdon, chairman of the Wellcome/CRC Institute for Cancer and Developmental Biology in Cambridge and master of Cambridge University's Magdalene College

Martin Bobrow, professor of medical genetics at Addenbrooke's Hospital in Cambridge

From 1 October they will be joined by:

Adrian Bird, director of the Wellcome Trust Centre for Cell Biology in Edinburgh

Jean Thomas, professor of macromolecular biochemistry at Cambridge University

Mark Walport, head of the Division of Medicine at Imperial College, London

Leatherback's survival will depend on an international effort

Sir—Until recently, marine systems have experienced remarkably few extinctions¹. But this state of affairs is changing, as reported in your News story “Researchers take US government to court over threat to turtles”² and by Spotila *et al.* in their Brief Communication³ on the potential extinction of the Costa Rican nesting population of leatherback sea turtles (*Dermochelys coriacea*). Many other long-lived marine organisms are also threatened, often because of unintended capture in fishing gear^{4–6}. A global solution is urgently needed to protect these species. It is possible: numbers of Kemp’s Ridley (*Lepidochelys kempi*), formerly the most endangered sea turtle, have increased tenfold since 1986 thanks to an international effort.

Leatherbacks nest primarily in four Pacific rookeries, and have declined dramatically over the past 20 years to about 250 in Mexico, 117 in Costa Rica, 2 in Malaysia and fewer than 550 in Indonesia. The Pacific Ocean may now contain as few as 2,300 adult females, making Pacific leatherbacks the world’s most endangered sea turtle.

Last autumn, a US judge closed one million square miles of the North Pacific to the Hawaii-based longline fishery, and the US National Marine Fisheries Service recently proposed another closure that would reduce leatherback takes by 45 per cent. But given the critical status of leatherbacks, this response is inadequate. Environmentalists believe that the United States must take decisive action, but closures in US fisheries alone will not resolve the problem—leatherbacks in both the North and South Pacific are killed by fishing vessels from several countries.

The leatherbacks’ nesting habitat must be protected throughout the Pacific. Direct harvest of eggs or turtles must be banned everywhere. Wherever fishing occurs, the bycatch of leatherbacks must be reduced to as close as possible to zero. Once countries such as the United States have minimized their own bycatch, they should encourage other nations to adopt protection measures. The US government and environmental organizations need to support research and management programmes in less developed countries, with environmental organizations educating the public and encouraging consumers to avoid products resulting from longline fishing.

Current treaties and organizations need to address the issue of leatherback bycatch

directly. The Convention on Conservation and Management of Highly Migratory Fish Stocks is at the final negotiation stages, and contains provisions to reduce bycatch; sea turtles should be explicitly included. The international agencies responsible for administering other conventions, including the Biodiversity Convention and the UN Convention on the Law of the Sea, should evaluate bycatch of protected species and report possible solutions. CITES (the Convention on International Trade in Endangered Species) has enforcement powers, but does not protect endangered species caught as bycatch.

The international community should prohibit trade in items caught in ways that harm endangered species—whether through an amendment to CITES or in a separate agreement. Whatever actions we decide to take, we must act quickly, or the 100-million-year reign of leatherback sea turtles will end in just a few decades.

Larry Crowder

Duke University Marine Laboratory, Nicholas School of the Environment, 135 Duke Marine Lab Road, Beaufort, North Carolina 29516, USA

1. Carlton *et al.* *Annual Review of Ecology and Systematics* **30**, 515–538 (1999).
2. Loder, N. *Nature* **405**, 495 (2000).
3. Spotila, J. *et al.* *Nature* **405**, 529–531 (2000).
4. Heppell, S., Crowder, L. & Menzel, T. in *Life in the Slow Lane: Ecology and Conservation of Long-lived Marine Animals* (ed. Musick, J.) 137–148 (American Fisheries Society, Bethesda, 1999).
5. Brander, K. *Nature* **290**, 48–49 (1981).
6. Casey, J. & Myers, R. *Science* **281**, 690–692 (1998).

Biotech pioneers have duties as well as rights

Sir—Your News report “California targets GM-trial vandals with new legislation” (*Nature* **404**, 799; 2000) states that a committee of the California state assembly has approved a bill to create tough penalties for the destruction of transgenic research crops. This is very good news, and I hope the bill will be approved by the judiciary committee and the full assembly.

However, I think the bill as it stands is unfair. At the same time as protecting GM trials, some obligations should be enforced on those holding trials. Protesters would then be more likely to accept the bill.

First, the GM trial itself needs legal protection. Some activists are bound to oppose the trial and application of the new technology of transgenic plants, as they have done for other new technologies when first introduced, for example nuclear energy. Approved insect-resistant and herbicide-resistant transgenic cotton, maize and soybean have been planted in many regions of the world for some time, and have had no untoward effects.

However, some radical protesters oppose

all aspects of GM technology and want to kill these transgenic plants ‘in the bud’—even though they have been approved—and prevent their development and application. Hence we need to protect properly planned GM trials and study via reasonable laws.

Second, GM trials and applications need reasonable regulation and control. Although transgenic plants have many merits, there may also be unknown environmental and ecological risks (although, up to now, no evidence of risk has been found). Thus, GM trials must be monitored, transgenic products must be labelled before being released into the market, and potential risks must be estimated. It is only fair to make these obligations enforceable by law as well.

Bao-Hong Zhang

Cotton Research Institute, Chinese Academy of Agricultural Sciences, Anyang, Henan 455112, China (Currently at: 1632 West 6th St Apt F, Austin, Texas 78703, USA)

Don’t ignore good work that you have to look for

Sir—Steven Hyman’s view is reported in your News report “BioMed Central boosted by editorial board” as “If I cannot access a paper freely I just ignore it” (*Nature* **405**, 384; 2000). Science is built on a base of acknowledgement and citation, yet this ideal is increasingly at odds with a culture which reflects short attention spans and demands instant gratification.

There is much to applaud in the establishment of BioMed Central, with its promise of free Web access to scientific papers. However, literature that is not instantly available does not deserve invisibility. If we are to keep faith with the past contributions of scientists, use of electronic media must be but one tool of literature review until such time as all literature (not just examples of what has been recently published) is available on the Web.

In my own field, the contributions of J. B. S. Haldane and S. Wright carry great weight. It is not possible to access the original work electronically, but it is fundamental to many of the current ideas in human population and disease genetics.

I hope that none of the giants on whose shoulders present-day scientists stand was careless enough to publish in a relatively obscure and unavailable journal or monograph. The precept “credit where credit’s due” should not evolve into “credit only when it’s convenient”.

Jerry Lanchbury

Department of Rheumatology, Guy’s, King’s College and St Thomas’ School of Medicine, King’s College, London SE1 9RT, UK

Database could give children safer medicines

Sir—We agree with Maurer *et al.*, who in their Commentary¹ define databases as “science’s neglected legacy”. Many large, complex databases are available to today’s scientists; information is stored to be accessed later and hypotheses can be tested using this well of knowledge.

However, the quality and effectiveness of databases can be uneven, and many areas remain starved of attention for a long time.

An example of these neglected areas in clinical medicine is the risk/benefit evaluation of medical treatment for children. Although we highlighted the need for a registry of randomized, controlled clinical trials in children long ago², so far no international databases on paediatric trials have been implemented.

This issue has recently come to the fore both in Europe and the United States, where children are being given unlicensed or unlabelled drugs because no appropriate trials have been carried out to meet the requirements of regulatory agencies³.

It is crucial that any medicines given to children should be evaluated both for efficacy and for safety. Although randomized, controlled trials are the foundation of evidence-based medicine, they are often conducted in secret and stopped prematurely. In many cases, these trials do not end with publication of the information that has been gathered⁴.

Children are at even greater risk than adults from inadequately tested drugs. This long-standing problem puts children in an underprivileged position within medicine. There is a clear need for new efforts to be made to have drugs properly investigated if they are going to be given to children. The development of an international database of paediatric trials is one way of opening up the process and avoiding these problems.

As Maurer *et al.* have valuably reminded us, society cannot get full value for its investment in science unless anyone who wants existing data can actually get hold of them. In this respect, recent developments in interactive health-communication technologies could make an online database of paediatric randomized, controlled trials easily accessible to scientists and clinicians. It would give them access to data on planned and ongoing trials, would avoid duplication of studies and would allow parents to obtain information on clinical studies in which they might give consent for their children to take part.

Such a database might thus also bring about a number of cultural and social benefits, including the potential to give people a

more direct role in caring for their children.

Piero Impicciatore, Chiara Pandolfini, Maurizio Bonati

Laboratory for Mother and Child Health,
Istituto di Ricerche Farmacologiche Mario Negri,
Via Eritrea 62, 20157 Milan, Italy

1. Maurer, S. M. *et al. Nature* **405**, 117–120 (2000).

2. Bonati, M. *J. Pediatr.* **116**, 667 (1990).

3. Conroy, S. *et al. Br. Med. J.* **320**, 79–82 (2000).

4. Bonati, M. *et al. Lancet* **355**, 1625 (1999).

Showman barely blinked at a dose of nerve gas

Sir—Your News article on nerve-gas tests (*Nature* **404**, 428; 2000) and the accompanying photograph brought back to me vivid memories of one of the key participants, prominently plonking himself right in the middle of the front row, and quite unmistakable from the cut of his well-tailored double-breasted suit, the



suitably placed handkerchief, cufflinks, the lot. I refer of course to the sartorial splendour of the “showman of organic chemistry”, Bernard C. Saunders (left, centre), who

enlivened the organic chemistry degree course that I took at Cambridge during the 1950s.

Halfway through each beautifully delivered lecture, when perhaps even the attention of the most industrious student began to flag, his “And that reminds me...” signalled another of his stories.

One of these was a personal recollection of nerve-gas tests on himself. After leaving the test chamber he wrote notes on the experience, but it was rather dark so they turned on the lights — which didn’t help much as the volunteers’ pupils had contracted under the influence of the gas.

But this work had an even wider significance: a sort of beating chemical weapons into chemical ploughshares, because compounds such as diisopropylfluorophosphate (DFP) enabled the Sanger/Hartley/Milstein *et al.* school to use such compounds as suicide (mechanism based) inhibitors to isolate the active sites of enzymes such as the serine proteases.

Bernard really was a great facilitator. For example, on reading Frederick Sanger’s 1988 review “Sequences, sequences, and sequences” I discovered that he had also provided the first sample of fluorodinitrobenzene (FDNB) to Sanger to test as an N-terminal reagent. It thus became known as the Sanger reagent, crucial for sequencing

small peptides and a big step on the way to Sanger’s first Nobel prize in 1958.

Derek T. A. Lampert

Professor Emeritus, DOE Plant Research
Laboratory, Michigan State University, East
Lansing, Michigan 48824, USA

Japan may seek embryo cells from overseas

Sir—According to Robert Triendl’s News report (*Nature* **404**, 321; 2000), reimplantation of human clones will probably continue to be prohibited in Japan, whereas experimentation into establishing human embryonic stem (ES) cells may be allowed.

On 17 December 1999, the cloning subcommittee (chaired by Yoshio Okada) of the Council for Science and Technology’s bioethics panel concluded that human cloning should be restricted by law. A law being drafted this year will primarily cover the prohibition of nuclear transplantation of human somatic cells. However, on 6 March, the same subcommittee indicated that research to establish human ES cells should be allowed, subject to guidelines being developed.

As a consequence, there is increasing expectation of immediate progress in research into establishing human ES cells. In some universities, study protocols for related experiments have already been submitted to the appropriate ethics committees.

However, progress in experimentation using human embryos cannot be guaranteed unless some significant problems are resolved, including regulation of proprietary rights of intellectual possession and profits to be derived from the establishment of such cells; ownership by those involved in the experiment and/or the donor of fertilized eggs; and problems related to the donation of fertilized eggs.

Concerning the last problem in particular, donation of fertilized eggs to outside facilities, even with the donor’s consent, may be against the guidelines of the Japanese Gynecology and Obstetrics Society. In any event, the society does not allow the use of fertilized eggs for anything other than infertility treatment.

If such problems are left unsolved, experiments to establish ES cells would be practically impossible even if the Council for Science and Technology does develop guidelines. Research using human ES cell strains from overseas may represent the most realistic option likely in Japan.

Tohru Nakanishi

Department of Biochemistry, Okayama University
Dental School, 2-5-1 Shikata-cho, Okayama City
700-8525, Japan

The great Dane of Uraniborg

The story of how astute manipulation of the system facilitated world renown.

On Tycho's Island: Tycho Brahe and His Assistants, 1570–1601

by John Robert Christianson

Cambridge University Press: 2000. 452 pp.

£30, \$34.95

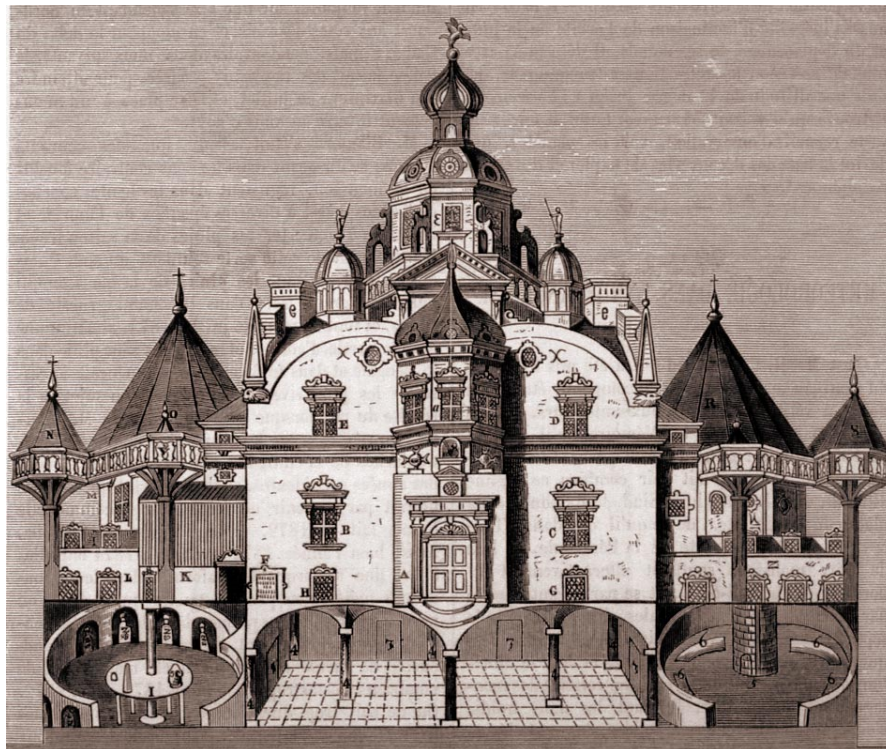
Kenneth J. Howell

Tycho Brahe (1546–1601) was arguably the greatest observational astronomer to live before telescopes were invented. His systematic observations of comets, planets and the supernova of 1572 launched him into world fame as a major figure of modern astronomy. In retrospect, it seems highly unlikely that such a monumental achievement could ever have been attained without Tycho's mastery of the patronage system and his managerial skills, which he consciously placed at the service of science. Christianson's book delightfully tells the story of how Brahe negotiated, built and maintained Uraniborg — one of the first modern observatories — on the island of Hven. It entertained a continuous stream of young assistants, many of whom subsequently became leading lights in the courtly, ecclesiastical and university circles of early seventeenth-century Scandinavia.

Christianson's work is the second of two major studies of Tycho's life to be published in English within the past decade which have thoroughly revised our understanding of him. The traditional image of Tycho as a hard-nosed empiricist, who offered a compromise system of the planets that lay between those of Ptolemy and Copernicus, turns out to be a rather superficial portrayal of the great Dane.

A more accurate picture had to wait almost four centuries, until the late Victor Thoren's 1991 biography, *The Lord of Uraniborg: A Biography of Tycho Brahe* (Cambridge University Press), revealed a far more complex thinker, whose combination of observation and theory paved the way for Johannes Kepler's theory of elliptical planetary orbits. Smaller historical studies unveiled a Tycho with wide-ranging knowledge not only of astronomy, but also of chemical experiments, Ovidian poetry, the Hermetic tradition and biblical interpretation. Tycho emerges as a Renaissance Neoplatonist who sought to unite astronomical, astrological and chemical information into a world picture that could not be reduced to any one of them. Christianson's book complements Thoren's by elucidating the web of social conditions, family ties, systems of support and royal favours that afforded Brahe unprecedented support for his research.

A central theme of the book is Tycho's 'familia', his highly structured and well-



Tycho's microcosm: the observatory at Uraniborg which Tycho ruled as a noble lord with royal support.

trained research community at Uraniborg which corporately engaged in big science. The author's colourful descriptions of familia dinners, evening entertainment and visits of famous guests, not only scientific, serve to create the ambience of life that surrounded diurnal duties in the laboratory below and nocturnal vigils in the observatory above. Tycho's familia drew on a pool of assistants from throughout Europe who sought to learn from his research, and also to share in the glory of his achievements. Although Tycho ruled Uraniborg as a noble lord with royal support — at times like a tyrant — the collection, recording and reduction of data often showed a remarkable reciprocity between master and assistants. In this way, Tycho's programmes of research on Hven and in Prague became harbingers of the collaborative nature of modern science.

What sixteenth-century social conditions did not allow was the separation of home and laboratory, so that partaking of the evening's delights and observing the stars followed upon one another as predictably as the phases of the Moon. Tycho's household attempted to model an ideal harmony of the world system in a microcosm of mundane work, social cohesion and the pursuit of knowledge.

This book mercifully refrains from focusing on power struggles within science, even

though Christianson occasionally hints at this as one of his main concerns. That Tycho was concerned with the social aspects of power there can be no doubt, but the attempt to explain scientific communities as engaged in power struggles has all too often been extended to the level of theories as purely social products. It is clear that Tycho's theories could not have been produced without the research support provided by the patronage system. It is equally clear that these theories were not determined by that patronage. Christianson's Tycho emerges as a man who used the complex web of patronage to pursue the pursuit of knowledge that would lead, at least in his own mind, to a unified system of truth that would benefit both science and humanity.

Part two of Christianson's book contains a catalogue of biographical sketches of Tycho's co-workers, providing information about their lives that has long lain hidden in obscure and inaccessible sources. The author's knowledge of the Scandinavian literature on Tycho brings to light many aspects of the astronomer's life that will be new to English readers.

As comprehensive as this second part appears, however, it does contain gaps that should be filled. For example, two important correspondents, Caspar Peucer and Henry Brucaeus, fail to appear in the catalogue



Meeting of minds: Tycho in discussion with Kepler at the Prague observatory.

because they never visited Uraniborg or Tycho's other research centres. Yet their exchange of letters with Brahe is essential for our understanding of late-sixteenth-century cosmological thought.

Despite such lacunae, Christianson's catalogue provides younger scholars with a treasure chest of nuggets they can use to expand our knowledge of early modern astronomy and its intricate social nexus. I have no doubt that Christianson's book will serve as a reference point for future historians of astronomy. ■

Kenneth J. Howell is in the Newman Foundation, University of Illinois, Champaign, Illinois 61820, USA.

More along these lines

Seeing and Believing: The Story of the Telescope, or How We Found Our Place in the Universe

by Richard Panek
Fourth Estate, £12

An alchemist of our times

Memory Effects

by Roald Hoffmann
Calhoun Press: 1999. 79 pp. \$9

Lars-Håkan Svensson

Poets who want, or need, to be more than "the unacknowledged legislators of the world", usually opt to be teachers, critics, editors, insurance executives, doctors or diplomats. Only a very few, however, have transgressed the critical border between the arts and science. One case in point is Miroslav Holub, who was a noted immunologist as well as a prominent poet. Another is Roald Hoffmann, who shared the 1981 Nobel Prize in Chemistry with Kenichi Fukui and who has since transformed himself into a poet and essayist. His most recent volume of poetry, *Memory Effects*, shows him to be an accomplished craftsman who moves effortlessly between the two cultures.

Although, technically, Hoffmann was a scientist before he was a poet, his artistic leanings are of long standing. A student of chemistry at Columbia University in the mid-50s, he almost switched to art history, and came under the spell of the distinguished literary critic Mark van Doren. At Cornell, where he has taught since 1965, he has long been a member of a small, informal group of writers which includes the poet A. R. Ammons. Although Hoffmann the poet is by no means a disciple of Ammons, it stands to reason that Ammons' peculiar combination of scientific exactness, close observation of nature, occasional colloquialisms and prophetic Romanticism must have made an impression on him.

From the outset, Hoffmann's poetry has oscillated between the world of science and that of personal experience, his subjects and language ranging from the general and the impersonal to the commonplace and the anecdotal. The reader is continually aware of

a nimble and curious intellect whose gift for abstraction is matched by a keen sensitivity to the concerns of everyday existence and a true poet's ability to connect one type of experience with another.

It is more than likely that the title of Hoffmann's first collection, *The Metamict State* (University of Central Florida Press, 1987), had some readers reaching for the dictionary; yet by the time the reader arrives at the final poem, where the amorphous, radioactive state in question is elucidated in no uncertain terms, it is clear that Hoffmann is not writing for a congress of chemists but for the alert common reader intent on expanding his or her range of experience. In fact, without falling prey to what literary critics refer to as the "pathetic fallacy", Hoffmann reveals how the world of science is replete with events and phenomena that, at least to the layman, parallel and illuminate those of ordinary life. The miracles — and disasters — he observes in his laboratory have counterparts in real life, not least in his own life.

Just what some of these are becomes more apparent as *The Metamict State* is followed by *Gaps and Verges*, his second volume of poetry (University of Central Florida Press, 1990), and then by *Memory Effects*. Most of the poems here are in Hoffmann's characteristic vein, revealing him as a shrewd observer of science, nature, himself and his fellow men. There is, for example, a poem that begins by pondering on the work of the Finnish chemist Olavi Erämettä, which then suddenly moves on to consider Väinöläinen, the Orphic hero of the Finnish national epic poem *Kalevala*. There is a poem on "The 1986 Nobel Prize in Physics", poems on nature, and poems that are meditations.

However, the strongest and by far the most personal part of the new collection is a sequence of poems relating in bare, precise language the unspeakable horrors that Hoffmann experienced in his childhood in Poland. These are poems of exceptional emotional intensity, stripped of all ornamentation, reflecting the poet's memories of escaping Nazi persecution by hiding in a school attic. One poem, perhaps the most moving of the entire collection, commemorates his father, who, after helping his wife and child escape, returned to the labour camp where they had all been held and was killed by the Nazis while organizing a break-out attempt.

Another poem describes the imaginary journey from Poland to California which son and mother fantasized about while in hiding. (After years of aimless itinerancy throughout Europe after the war, they were eventually able to emigrate to the United States in 1949.)

Beside such poems, even the most successful of the other poems might at first

New in paperback

The Big Bang: A History of Explosives

by G. I. Brown
Sutton Publishing, £9.99, \$16.95

The Truth of Science: Physical Theories and Reality

by Roger G. Newton
MIT Press, \$16.95, £11.50

A Passion for Birds: American Ornithology after Audubon

by Mark V. Barrow Jr

Princeton University Press, \$19.95, £12.50
"Barrow exposes a rich portrait of human strength, frailty and ego in the transformation of a popular nineteenth-century hobby into first a scientific discipline and ultimately an influential mainstream profession... I was gripped from cover to cover by this book."
John W. Fitzpatrick, *Nature* 397, 119 (1999)

Insights of Genius: Imagery and Creativity in Science and Art

by Arthur I. Miller
MIT Press, \$18.95, £12.50

appear insignificant. However, they contribute to the overall effect of Hoffmann's third collection as a vision of the world that encompasses both suffering and wonder, anger and beauty, the incidental and the recondite. Hoffmann's blending of such seemingly incompatible elements and experiences is a considerable achievement. His book is a memento both to those who think that science is devoid of poetry and to those who doubt poetry's ability to express pain and loss in the modern world. ■

Lars-Håkan Svensson is in the Department of English, Helgonabacken 14, Lund University, 2224 67 Lund, Sweden.

Sex, leaves and rock and roll

The Ingenious Mr Fairchild: The Forgotten Father of the Flower Garden

by Michael Leapman

Headline: 2000. 280 pp. £12.99

Sandra Knapp

A few years ago a friend told me that gardening was the new rock and roll. I laughed at the time, but since then I have seen those very words in Sunday supplements and have heard them from a wide variety of people. Gardening has become enormously trendy, cutting across all parts of society, from lords and stockbrokers with acres of land to those of us with window boxes in tiny flats. Anyone near Vincent Square in central London when the Royal Horticultural Society holds its monthly shows might argue that plants have never been more popular. Attendance at the famed Chelsea Flower Show has increased so much that it is difficult now to see the plants for the crowds.

Thomas

Fairchild, author of *The City Gardener* — a 1722 treatise on how and what to grow in smoky eighteenth-century London — would be pleased. He saw gardening as an occupation for everyone, and put this into practice by cultivating and ultimately creating plants that thrived in trying circumstances. Michael Leapman's book is ostensibly about Thomas Fairchild, but is really an entertaining skate over the surface of gardening history,

using Fairchild's creation of an interspecific hybrid as a departure. His story goes in two directions, to the past when early European botanists discovered the sex life of plants, and to the future, where genetic modification appears to threaten the very harmony of nature.

The early eighteenth century was a time of great intellectual excitement in Europe, of which Londoners of the time considered their city an integral part. Patrons of science — as collectors of objects — are given somewhat short shrift by Leapman, although he does acknowledge their contribution to future scientists. He seemingly considers gardeners such as Fairchild, who worked with living things, to be more scientific. I feel this does a disservice to men such as George Clifford in Holland and Sir Hans Sloane in London. Without Sloane's herbarium we would have no record at all of "Fairchild's mule", the hybrid of a sweet William (*Dianthus barbatus*) and a gillyflower (*Dianthus caryophyllus*), except through mentions in letters and the writings of Fairchild's contemporaries. Without this specimen the book might well have been called "The Madly Entertaining Mr Bradley", because much of the documentary evidence about Fairchild appears to have been derived from the writings of

Richard Bradley, whose career spanned all the eccentricity of the age. In a way I wish Leapman had written a book about him — the bits about Bradley in this book made me laugh aloud. Fairchild's story is reconstructed, on the basis of quite scanty evidence, as a metaphor for society's fear of the new.

That plants engaged in sexual intercourse was the butt of many Georgian (and later) jokes and poems, several of which are contained in the book. Erasmus Darwin, grandfather of Charles, penned a verse referring directly to Fairchild's mule — or so it seems — in 1781, long after Fairchild's death. Linnaeus's sexual system, which he used for the first time in 1735 in his *Systema Naturae*, allowed botanists to organize the plant world into categories based upon observation, rather than upon prior knowledge of their medicinal properties. This, like gardening, was a great leveller. Just by counting reproductive parts, anyone could look at a flower and try to figure out what it was. Thus began the democratization of science.

Leapman assumes that the botanists did not experiment with hybridization, that only gardeners did that, but the evidence is just not there. We know about Fairchild's mule because a specimen exists, not necessarily

One man's weed ...

... is another man's flower. This is the gist of perhaps the least pretentious of the mercifully few aphorisms in this otherwise beautiful book, *Weeds* (Chronicle Books, \$19.95, £14.99). Chicago-based Howard Bjornson's lavish collection is a

photographic portraiture of the American weed world, or, as the book puts it, of the "free spirits of the natural world". Most of the weeds featured in this book were gathered within a few miles of Bjornson's studio.



because it was the first hybrid created. This in no way detracts from Fairchild's contribution: for his work in transferring pollen from one species to the stigma of another, raising the resulting offspring and distributing them among his customers, he can indeed be considered "the father of the flower garden". Whether his creation was looked upon with horror by his contemporaries is not clear. But in the long run it cannot have been, because the flower spread, and hybridization to produce new garden varieties became quite commonplace.

It's not really a surprise that hybridization was shocking to the religious sensibilities of the eighteenth century, but to contend that Fairchild's legacy for an annual sermon — on "The Wonderful World of God in the Creation or on the Certainty of the Resurrection of the Dead proved by Certain Change of the Animal and Vegetable Parts of the Creation" — was given to assuage his guilt in creating an interspecific hybrid is going perhaps a bit far.

Leapman's book concludes with a short foray into the genetic-modification debate. His discussion is thought-provoking, and by reminding us that modification exists in the garden as well as the farm, it reveals a somewhat alarming side of the current public debate. The question, "who needs a blue rose?" would not have occurred to Fairchild and his contemporaries, who — although deeply religious and believers in God's design — were intensely curious about the world around them.

Despite a hard-headed, ostensibly agnostic and seemingly pragmatic outlook on life, people still seem to yearn after an 'organizer' or a higher 'intelligence' who has set up nature in a fixed design with which we meddle at our peril. It is an easy option to leave it all to a designer, but much harder to examine evidence impartially and rationally and to think carefully about alternatives. The easy option can lead to a frightening fundamentalism, whereas the second is the responsible route. Sentimentality about nature and its products would have seemed very strange to Thomas Fairchild.

Sandra Knapp is in the Department of Botany, The Natural History Museum, Cromwell Road, London SW7 5BD, UK.

More from the garden

The Tulip

by Anna Pavord
Bloomsbury, £30

Millions of Monarchs, Bunches of Beetles: How Bugs Find Strength in Numbers

by Gilbert Waldbauer
Harvard University Press, \$24.95, £15.50

Science in culture

Collisions and encounters

Science, representation and art at CERN.

Martin Kemp

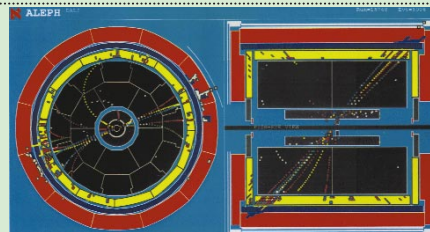
Hugely complex issues of representation lie at the heart of communication in modern physics, whether person-to-person or at the interface of humans and machines. It is potentially exciting, therefore, that a group of European artists are currently involved with scientists at CERN, the laboratory for elementary particle research in Geneva. The project, "Signatures of the Invisible", moved into gear at the start of this year under the direction of the film-maker Ken McMullen, as a collaborative venture of the London Institute (a conglomerate of colleges of art and design) and physicists at CERN. During this year, 11 selected artists will spend periods at Geneva in dialogue with scientific collaborators.

Whatever visions and skills the individual artists may bring to the project, they will be hard-pressed to match the optical, perceptual and geometrical sophistication of the programme DALI, devised by the CERN physicist Hans Drevermann for the depiction of the reactions between elementary particles in the ALEPH detector. The question he has posed himself, in the face of the vastly increased complexity of events recorded in the CERN detectors, is whether "a fast, efficient and unambiguous transfer of data to the human brain via visual techniques [is] still possible for complicated events". The task he has set is to adapt various techniques of two- and three-dimensional rendering and methods of schematic and abstract projection in combination with diagrams to provide "new visual representations" that are "better tuned to the capabilities of human perception" than existing methods.

Broadly speaking, four types of representation are potentially available. The most obvious and traditional — depending ultimately on the Renaissance methods of rendering objects systematically in space — use techniques of 3D projection and rotation to convey an intuitively clear picture of the detector as if cut away to reveal the events within. The second type of projection is derived from perspective projections, exploiting the characteristic that the size of detector and event unit are decreased the further they are from the centre (the so-called 'fish-eye' projection; see figure).

The third representation exploits the cylindrical geometry of the detector and event, leading to intuitively understandable projections resembling a cut through the cylinder, or to more abstract schemes. These may involve the unrolling of the cylinder or the deforming of a wedge from the circle into a rectangle. Such methods aim to enhance aspects of comprehension in selective ways that schematize normal visual appearance yet still play to our customary modes of perception and understanding.

The fourth representation resembles the kind



'Fish-eye' representation of an ALEPH event.

of map projection devised by Mercator. This method involves technical abstraction in which aspects of the data are plotted visually in arrays that are analytically powerful but essentially unsuitable for presentation to non-specialists, as they cannot be intuitively understood through everyday acts of perception and cognition.

Such acts of picturing, if they are to be as effective as possible in conveying complex events, operate at the limits of visual legibility, as point size and linewidth must be as small as possible to cope with the large information content of the events. This, in turn, drastically affects human colour perception, and the choice of foreground and background colours thus becomes crucial.

Computerized pattern recognition and analysis programmes achieve results that go beyond what human vision can accomplish, yet Drevermann is firmly committed to methods that will enable our senses to compete when utilizing the minimum point size and linewidth that can be discerned in various media. Why bother, if the ability of computers to read data and interpret numbers is so powerful? First, because communication between scientists is very effective using visual means; these means become even more desirable if scientists are to communicate with a lay audience — of huge significance, given the need for continued public funding. Second, because even the most powerful computers and programmes need to be validated by the scientist, using the most powerful computational device of all, the human brain.

The blending of intuitive perception of the visual with intellectual analysis of the data provides a vital tool of verification. It is at this point of blending that certain processes of understanding may mysteriously occur, often unbidden. Means of communication and modes of understanding operate by symbiosis. As Drevermann has said: "It was not my initial goal to generate pictures of large publicity value nor pictures that were appreciated by artists. This came as a surprise. However, it seems that the scientific value of the pictures has improved, since I am aware of their non-scientific value."

From this starting point, it will be fascinating to see how the encounters at CERN between art and science will bear fruit. (See <http://alephwww.cern.ch/DALI/>.)

Martin Kemp is in the Department of the History of Art, University of Oxford, 59 George Street, Oxford OX1 2BE, UK.

The well-spring

About 3,000 years ago, the Greeks invented science.

Lewis Wolpert

In reflecting on and enjoying the scientific triumphs of the past millennium, we should also honour the origins of science and give grateful thanks to the Greek scientists of the millennium before Christ. There is no earlier society in which one can identify science as distinct from technology, and where there was an objective attempt to explain the way the world works. The Chinese — who had a magical view of the Universe — were quite different.

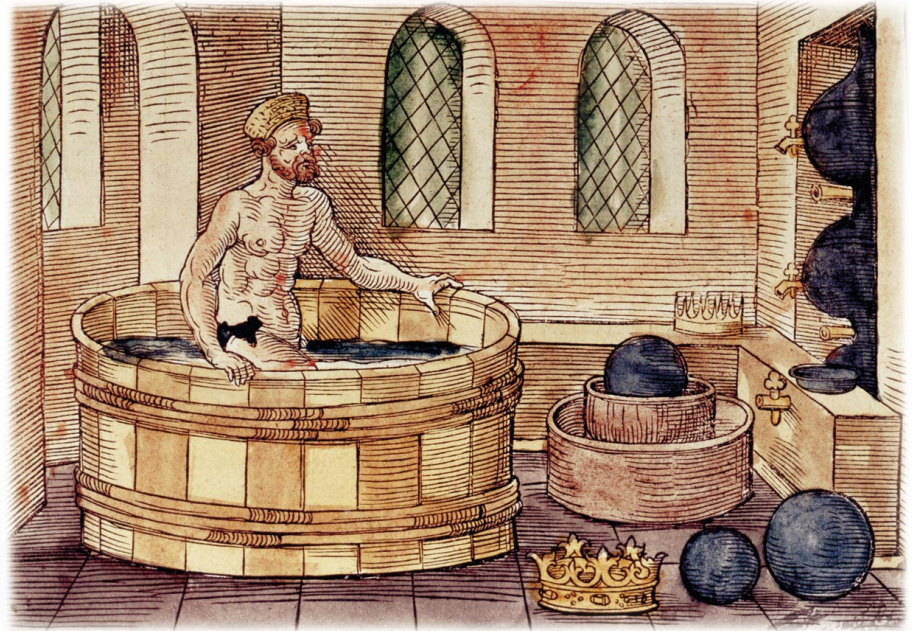
In particular, it was the Greeks of Ionia who were the first to attempt to explain the world in concrete terms as distinct from mystical ones. They held the belief that there were general laws that could be discovered, and the conviction that rational argument was essential.

Science is not a natural mode of thought, as the world is not built on the expectation that we gain from our everyday experiences. The Greeks stood back from nature and tried to understand it for its own sake; indeed, as the historian of Greek science Geoffrey Lloyd has suggested, they may even be thought of as having invented the idea of nature. Understanding was to be its own reward.

This process is beautifully illustrated by the first record of a scientific theory, that of Thales of Miletos. In about 600 BC, Thales suggested that everything was made of water in different forms. Water could change its form from solid to liquid and back again, and water was essential for life. This is a fantastical suggestion, against all common sense, but clearly a scientific theory that could be tested. The possibility of objective and critical thinking — and, crucially, open debate — about nature had begun. Anaximander, for example, strongly disputed Thales' claim about water and instead proposed air as the key substance.

Although Thales has the honour of being the first person to do real science, we must also recognize that he was almost certainly aware of earlier achievements in mathematics, particularly those of the Egyptians and Babylonians. Yet it was Thales who first made formal mathematical statements such as: a circle is bisected by its diameter. And: if two straight lines intersect, the opposite angles are equal. Hence he laid the foundations for geometry and Euclid.

My hero among the Greeks is Archimedes, who followed in the footsteps of Aristotle and Euclid by stating postulates and then deducing the logical and formal consequences. In mechanics, he invented the



To those who insist that scientific knowledge is transitory, Archimedes' work is an elegant rebuttal. He is the first mathematical physicist and applied mathematician, and no one else made any progress in his area for another 1,500 years.

concept of the centre of gravity and provided a geometrical proof that two bodies of different sizes placed on a simple balance will be in equilibrium at distances reciprocally equal to their weights. Even today, it is non-trivial for a physics student to arrive at such a proof. Archimedes also created hydrostatics. Just consider the achievement of his second postulate: "Let it be granted that bodies which are forced upwards in a fluid are forced upwards along the perpendicular to the surface which passes through their centre of gravity."

From such postulates Archimedes showed that a body's loss of weight in a fluid is equal to the weight of water that it displaces, and went on to discover the specific gravity of substances. To those who insist that scientific knowledge is transitory and continually replaced, his work is an elegant rebuttal. He is the first true mathematical physicist and applied mathematician, and no one else made any progress in his area for another one and a half millennia. No wonder Galileo called him "divine".

The Greeks had a society in which there

was vigorous debate and discussion of evidence. It is also the first society where an individual author explicitly distances himself from the received tradition and criticizes it, and even claims originality for himself. The admiration of one's peers is one of the major rewards of science; in Greece this practice first became possible, as authors adopted the first-person singular for the first time in history. Perhaps this had its origin in the demand for recognition by the Greek poets and the Greek tradition in examining evidence in the context of law and politics. The success of science depends on our having inherited this openness and the right to challenge authority. This truth should be forcibly pointed out to those who wish to constrain scientific investigations. ■

Lewis Wolpert is in the Department of Anatomy and Developmental Biology, University College London, Gower Street, London WC1E 6BT, UK.



In the days of the comet

Dispatch 135 from the Oort Cloud Survey.

John M. Ford

Camfield is dead, and this ship is very quiet now. I have tried to be hopeful in the recent dispatches: we were, Camfield certainly was. Prions are not supposed to kill people any more, but they can, and they have. Which is part of the reason Camfield was out here in the first place.

He was a teller of jokes and he played the guitar very well — these are valuable things when you are doomed to spend years aboard a cantankerous old ship. Several instalments ago, I described the lab accident that infected Camfield, and I have received numerous messages calling the events absurd. This is true. In addition to myself, the organic Petrovna and the Neumann Thucydides saw the incident, and we all laughed until we realized Camfield was hurt. Petrovna, at least, can forget, although I do not think she will.

The prion has been decrypted and entered into the antigenic database, so no one should ever again die of Agent Op-1175s/CFD.

Which is the story, but not its point.

At the cusp of this millennium we discovered that it was not hard to manufacture prions, and not that hard to custom-twist them. It took longer for our twists to be meaningful, but now organic humanity can don an armour of proteins for defence against a hostile Universe. Rather like viruses. Draw your own conclusions.

If one could find the right message, a prion would make a wonderful interstellar, even intergalactic, postal card: immune to temperature, pressure, radiation and time. The ideal pony for the express would be a comet, packed with messenger proteins, flung into a hyperbolic orbit, to seed any worlds at the far end with its cargo.

One could write one's name in the evolving life of a planet. At exactly the right moment, one might even begin the process, dropping a bouillon cube into the primordial soup.

Assuming that no one at the other end is quite as evolved, and quite as dependent on delicate higher neural functions, as we are.

So here we are, myself, 29 (down from 30)

organic crew, and eight Neumänner, combating the comets of the Oort for prions. We have found a lot of prions, and there are a lot of comets left. You've got mail, as we said when I was organic.

Maybe. Or maybe one of the 48 published theories of spontaneous prion formation in comets is correct. It is the Neumänner who are most insistent on deliberate seeding. Perhaps it comforts them to think that, just as we built them, somebody built us. How human of them — but, as their namesake said, adequately describe any activity, and a machine can perform it.

In Camfield's last hours he was afire with fever, his whole body trembling, but there was a clarity in his speech that was at once heartbreaking and terrifying. Fischer, Chiang, and the Neumann Hypatia were tending him. Abruptly he calmed, fixed Chiang (and me, unavoidably) with a direct stare, and said, "I see the Martians now! They are flat, and they roll!" He shivered then, and I heard his heart stop.

The exclamation points are not added for drama. He was excited by what he saw, transported by whatever the alien messenger in his brain was revealing to him. Camfield was born on the Moon, not Mars, so we cannot explain away the vision as *Heimsucht*.

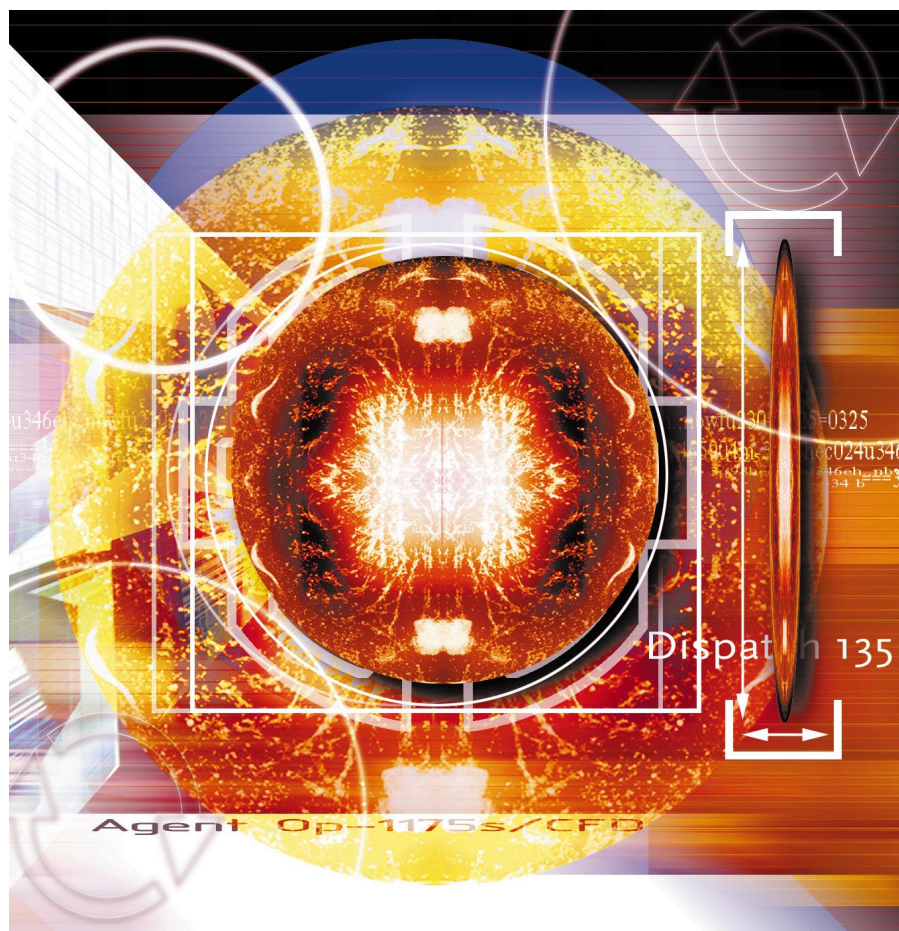
We cannot, of course, positively explain it at all. But we must examine the possibility that, eons ago, Op-1175s/CFD fell on Mars and began life there, which was later carried to Earth by a planetary blunt trauma.

Thucydides carefully wrapped and sealed Camfield's remains for storage until we return to the Moon, eight years from now. When he was done, Sid paused for two full minutes (exactly — we are like that), just looking at the bundle.

This kind of behaviour is by no means strange in a Neumann (one can adequately describe a thoughtful pause) but I asked Sid what he was thinking. He waited fourteen seconds longer — which was purely theatrical of him — and said, "I will miss Camfield. He was always interesting to be with, even when nothing was said. And has he not left us with a fine and difficult question?"

Camfield gave many gifts to his shipmates and his ship. The question — and it is fine — he gave to all of us.

Autonomous Exploration Vessel John M. Ford was one of the first private citizens to undergo cortical shift, first to a netlinked mainframe, passing through a series of mobile installations to an AEV hull. In addition to his fiction (such as The Last Hot Time, to be published by Tor books) he is at work on The League of Steersmen, a history of the AEVs.



Neural circuits in silicon

Chris Diorio and Rajesh P. N. Rao

Studies of neurally inspired silicon circuits are showing how networks of neurons can select and multiply input signals. They may also provide alternative ways to build computers modelled on biology.

Animal brains form the centrepiece of nature's development of information-processing machines. In attributes such as adaptability and fault tolerance they are likely to remain unsurpassed by any human-built machine in the foreseeable future. Digital computers form the centrepiece of the past quarter-century's explosion in information technology. In terms of numerical computation, their capabilities far exceed those of animal brains. This dichotomy perplexes us. How do animal brains compute? And if silicon technology is so powerful, why can we not build thinking machines?

Carver Mead, a pioneer in integrated-circuit technology, quipped during a discussion of brains versus computers that "silicon doesn't know anything about bits". The implication is that nothing about silicon itself requires computers to be digital. On page 947 of this issue, Hahnloser *et al.*¹ take this reasoning to its logical conclusion: they have built a cortex-inspired silicon circuit that multiplies and selects features in its input, using a network of neuron-like elements. The theoretical basis for this work is not new: in 1996, Salinas and Abbott² reported computer simulations of a network of model neurons, demonstrating that the network could perform these exact tasks. What Hahnloser *et al.* have done is to reproduce the behaviour of such a network in silicon.

The networks studied by Salinas and Abbott and by Hahnloser *et al.* both use recurrent (or feedback) connections that are excitatory for connections between neighbouring neurons and inhibitory at larger distances (Fig. 1a). This pattern of local excitation and long-range inhibition is common in contemporary models of the cerebral cortex^{3–6}. One attribute of this type of network is that, when a neuron at a given physical location receives an input, the network responds by activating both the stimulated neuron and a cluster of neurons around it (Fig. 1b). Moreover, when the background input to all neurons is increased systematically, this cluster of activity is multiplied by a gain factor that is a linear function of the background input (Fig. 1b). These responses are intriguing from a neurophysiological perspective. Say, for example, we equate the background input to a motor signal representing eye position, then these multiplicative responses are similar to the modulation of neuronal

responses by eye position observed in the visual cortex⁷.

A second important feature of this type of recurrent network is its ability to select

a single stimulus when presented with several competing stimuli. In this case the network exhibits nonlinear behaviour, selecting the strongest of the competing stimuli and

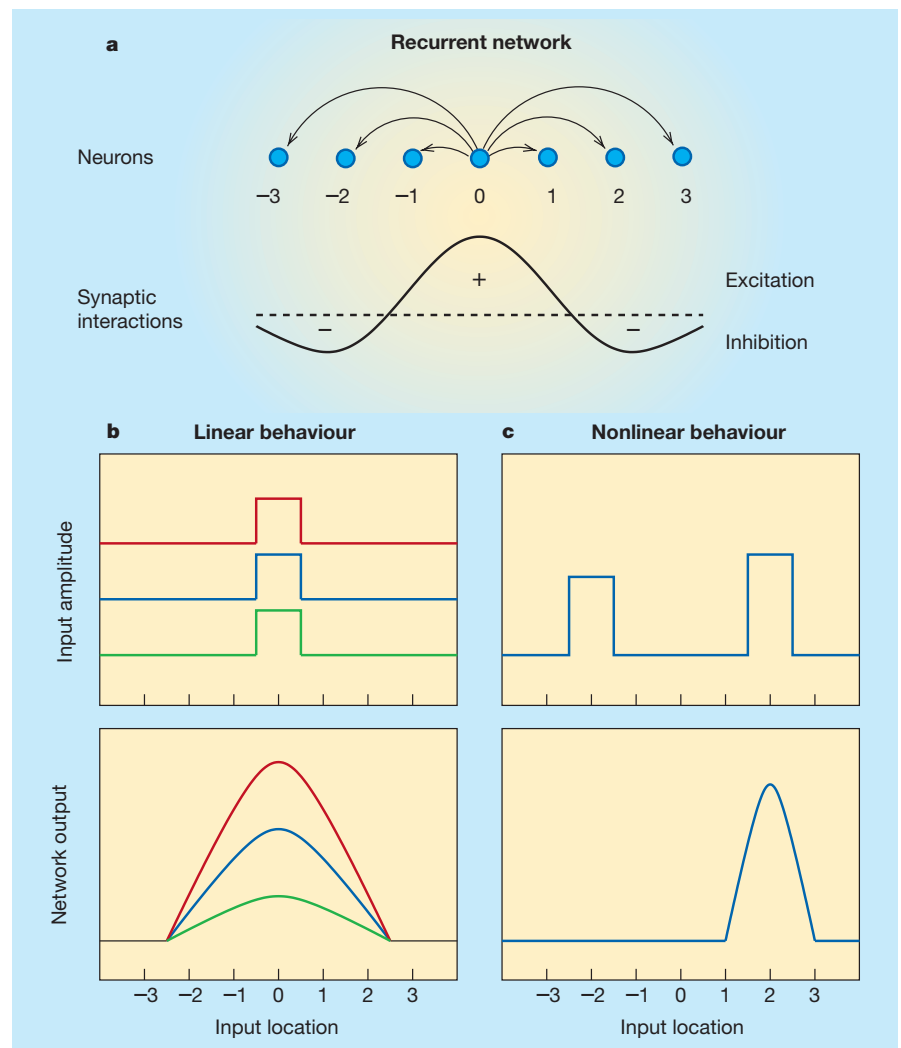


Figure 1 Linear and nonlinear behaviour in a silicon circuit inspired by neurobiology.

a, The neuronal network studied by Salinas and Abbott², and built in silicon by Hahnloser *et al.*¹.

Each neuron excites its nearest neighbours, and inhibits neurons further away from it. The artificial neurons, in both cases, use a continuously varying signal to model the firing rate of a cortical neuron. In their numerical simulations, Salinas and Abbott model the strength of synaptic connections from a neuron to its neighbours as a 'Mexican hat' function (shown below the network), whereas Hahnloser *et al.* approximate this function in their silicon circuit. b, Example of multiplicative responses (that is, linear gain modulation) in the network. Given three input pulses at the same location but at different background amplitudes (green, blue and red), the network multiplies the output by a gain factor that is a linear function of the background input. c, Example of stimulus selection in the network. Given input pulses at two different locations, the network selects the location with the stronger input, suppressing the location with the weaker input. This competitive, nonlinear behaviour is the result of recurrent excitatory and inhibitory interactions among neurons.

suppressing the weaker ones (Fig. 1c). Such behaviour can be regarded as a simplified form of sensory attention, whereby the network selects the stimulus location based on stimulus strength. From a neurophysiological perspective, the selection of a single target and the suppression of distractors is important, for example, in programming arm or eye movements.

The network's ability to switch between linear and nonlinear behaviour, based on its input, is very different from that of standard electronics. Engineers usually require separate analogue and digital circuits to carry out linear amplification and nonlinear selection, respectively. Hahnloser *et al.* explain their hybrid analogue–digital circuit in terms of the set of neurons that are active in steady state, and a gain that depends only on the identities of the active neurons and not on their analogue responses. Behaviour that derives from a common set of active neurons is linear in the input, whereas behaviour that derives from a comparison among different sets of active neurons is nonlinear in the input.

Hahnloser *et al.* extend previous studies of this class of recurrent networks by analysing the stability of the dynamical equations that model the network, and show that there are inviolate constraints on the allowed network states. In particular, they show that if no single input can activate a set of neurons (the set cannot form a memory), then no input can activate a supergroup of these same neurons (no supergroup can form a memory). A limitation of the study, however, is the requirement that synaptic connections be symmetric — that is, the strength of the connection from neuron A to neuron B must be the same as that from B to A. This assumption allows them to prove network stability, but is difficult to justify neurobiologically. A fruitful area for future research, therefore, is investigation of recurrent cortical networks with non-symmetric synaptic connections (see, for example, refs 8–10 and Supplementary Information¹).

What do Hahnloser *et al.*, and others like them, hope to accomplish by building silicon circuits modelled on biology? First, they can learn how to map neuronal 'primitives' (such as neurons and synapses) onto silicon, and then how to compute using these primitives (see, for example, refs 10–12). Second, they can investigate how physical and technological limits, such as wire density, signal delays and noise, constrain neuronal computation. And third, they can learn about alternative models of computation. Biology provides examples of non-digital computing machines that are incredibly space- and energy-efficient, and that excel at finding good solutions to ill-posed problems. Scientists may eventually decipher all of nature's electrochemical circuits, but the work of Hahnloser *et al.* demonstrates that we

already know enough to begin building integrated circuits that compute like biology. ■

Chris Diorio is in the Department of Computer Science and Engineering, University of Washington, Box 352350, 114 Sieg Hall, Seattle, Washington 98195-2350, USA.

e-mail: diorio@cs.washington.edu

Rajesh P. N. Rao is at the Sloan Center for Theoretical Neurobiology and the Computational Neurobiology Laboratory, Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, California 92037, USA.

e-mail: rao@salk.edu

1. Hahnloser, R. H. R., Sarpeshkar, R., Mahowald, M. A., Douglas, R. J. & Seung, H. S. *Nature* **405**, 947–951 (2000).
2. Salinas, E. & Abbott, L. F. *Proc. Natl Acad. Sci. USA* **93**, 11956–11961 (1996).

3. Ben-Yishai, R., Bar-Or, R. L. & Sompolinsky, H. *Proc. Natl Acad. Sci. USA* **92**, 3844–3848 (1995).
4. Somers, D. C., Nelson, S. B. & Sur, M. *J. Neurosci.* **8**, 5448–5465 (1995).
5. Douglas, R. J. *et al. Science* **269**, 981–985 (1995).
6. Stemmler, M., Usher, M. & Niebur, E. *Science* **269**, 1877–1880 (1995).
7. Andersen, R. A., Essick, G. K. & Siegel, R. M. *Science* **230**, 456–458 (1985).
8. Abbott, L. F. & Blum, K. I. *Cereb. Cortex* **6**, 406–416 (1996).
9. Rao, R. P. N. & Sejnowski, T. J. in *Advances in Neural Information Processing Systems* Vol. 12 (eds Solla, S. A., Leen, T. K. & Müller, K.-R.) 164–170 (MIT Press, Cambridge, MA, 2000).
10. Westerman, W. C., Northmore, D. P. M. & Elias, J. G. *Analog Integrated Circuits Signal Process.* **18**, 141–152 (1999).
11. Diorio, C., Hasler, P., Minch, B. A. & Mead, C. *IEEE Trans. Electron Devices* **44**, 2281–2289 (1997).
12. Mead, C. *Analog VLSI and Neural Systems* (Addison-Wesley, Reading, MA, 1989).

Neurobiology

Self-repair in the brain

Anders Björklund and Olle Lindvall

In tissues that can repair themselves, such as skin and liver, dead cells can be replaced either by the proliferation of nearby cells or by the activation of resident stem cells — undifferentiated cells with the potential to generate many different cell types. The brain apparently lacks this regenerative capacity, making it particularly vulnerable to injury or disease. Cells with stem-cell-like properties are likely to occur throughout the adult central nervous system, but they normally give rise to neurons in only a few, restricted areas. These cells might represent a dormant capacity for neuronal repair, if they could be mobilized to generate new functional neurons in response to injury. On page 951 of this issue, Magavi, Leavitt and Macklis¹ report the first intriguing evidence that this may be possible.

In the adult brain, the generation of new neurons (neurogenesis) occurs in just two regions². The first is the subventricular zone (SVZ) in the wall of the lateral ventricle. Here, new interneurons are generated for the olfactory bulb (Fig. 1a), which is involved in sensing odours. The second is the subgranular zone of the dentate gyrus, which gives rise to a different type of neuron, the granule cell. In these areas, there is seemingly a continuous turnover of interneurons and granule cells, implying that the newborn neurons replace dying cells. But does neurogenesis actually help to replenish damaged neuronal circuits? There is some evidence that damage to granule cells can trigger the increased proliferation and recruitment of new granule cells from resident progenitors³. And insults — such as seizures and inadequate blood supply to the cerebrum — that cause cells in the hippocampus to die are accompanied by increased neurogenesis in the dentate subgranular zone^{4–6}. But as yet there is no direct

evidence that the degenerated neurons are actually replaced by the new ones.

In other parts of the central nervous system, neural progenitors contribute to the ongoing formation of new non-neuronal cells called astrocytes and oligodendrocytes, and may participate in the reaction of these so-called glial cells to injury, as well as in scar formation^{7,8}. Other progenitors appear to remain undifferentiated, or die^{7,8}. A low level of neurogenesis has been described in parts of the intact neocortex of adult monkeys⁹, but the newly generated neuron-like cells appear to survive only transiently, and may not differentiate into fully functional neurons¹⁰.

In line with previous findings^{11,12}, Magavi *et al.*¹ have observed the continuous formation of new cells in the intact neocortex of the adult mouse. These proliferating cells were distinguished by labelling them with bromodeoxyuridine, a thymidine-base analogue that is incorporated into the new DNA formed during DNA replication in dividing cells. None of these newly formed cells, however, expressed any marker proteins characteristic of neurons.

To investigate the effects of neuronal death in the neocortex, Magavi *et al.* destroyed a subset of pyramidal neurons that project from the neocortex to the thalamus, which is a major subcortical relay station involved in the control of cortical function. Their technique resulted in the slow death by apoptosis of the targeted cells only, without affecting the surrounding cortical tissue. The number of new cells formed in the two weeks after the lesion was similar in control and experimental mice. However, about 1–2% of the newly formed cells in the damaged neocortex (about 50–100 cells per cubic millimetre) expressed neuronal markers.

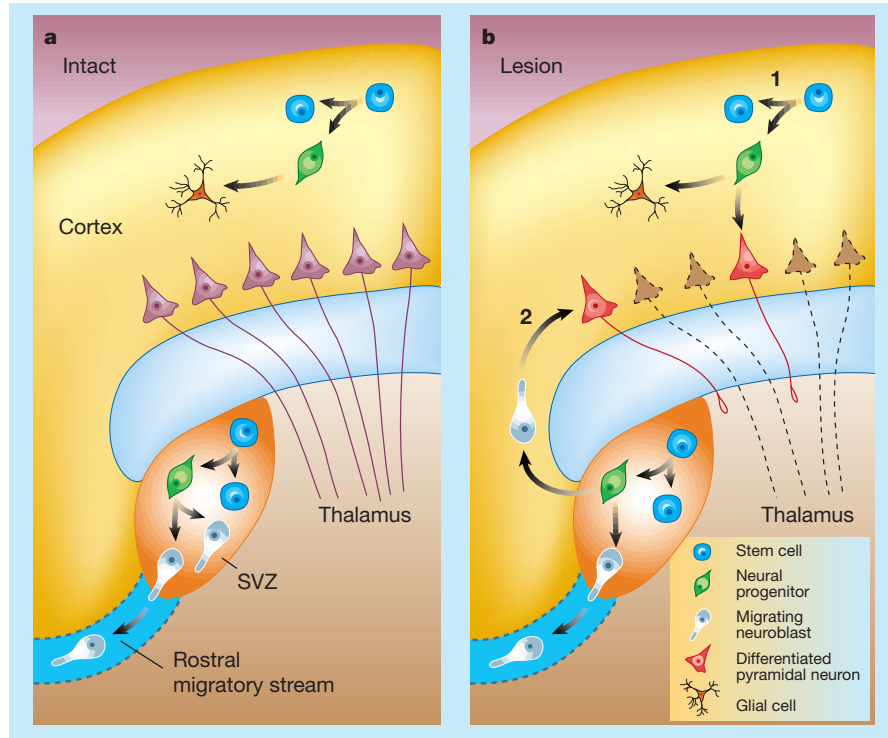


Figure 1 Does the brain have a dormant capacity for self-repair? **a**, In the neocortex of the adult mouse, stem cells and other neural progenitors occur locally in the cortical grey and white matter, and in the underlying subventricular zone (SVZ). The cortical progenitors give rise to glial cells (non-neuronal cells) only. The neurons generated in the SVZ migrate along the rostral migratory stream to the olfactory bulb (not shown). **b**, Magavi *et al.*¹ used a refined photolytic lesion to induce apoptotic cell death in a subset of neurons that project from the neocortex to the thalamus. The cells were filled from their terminals in the thalamus with nanospheres carrying a chromophore, which becomes toxic when irradiated with red laser light. During the following weeks, Magavi *et al.* detected new neurons with features of cortical pyramidal neurons in the layers undergoing degeneration. The new neurons appeared to be recruited from the resident cortical progenitors (pathway 1) and/or from the underlying SVZ (pathway 2). A few of these cells seemed to connect back to the thalamus, suggesting a hitherto unknown capacity for neuronal replacement and functional repair.

Such cells were not observed in the control mice. The new neuronal cells occurred only in the cortical layer that was undergoing degeneration, and some of them had a morphology characteristic of cortical pyramidal neurons. The newly formed neurons extended processes to the original target sites in the thalamus, suggesting that they joined up the damaged circuitry.

The observations are provocative, but how certain can we be that the interpretations are correct? As pointed out recently¹⁰, bromodeoxyuridine may label cells that are repairing damaged DNA rather than replicating DNA ready for cell division. So there is a risk that neurons undergoing cell death, as well as newly divided cells, may have incorporated the label. But bromodeoxyuridine-labelled cells in the damaged neocortex expressed markers — not found in controls — characteristic of migrating young neurons, and survived for at least six months. So it seems unlikely that they were dying. It is also conceivable that a few surviving neurons might be labelled as a result of reversible DNA damage. This was not controlled for, but in this case one would expect less

bromodeoxyuridine to be incorporated than was observed.

These intriguing data raise several questions. First, we need to know where the newborn neurons originate. Possibilities are the progenitors present locally in the cortex, or cells in the underlying SVZ region (Fig. 1b). During postnatal development, SVZ cells migrate to neocortex to form glial cells, but in the adult they migrate only to the olfactory bulb. Yet adult SVZ cells can respond to damage by increased proliferation and migration towards the lesion^{13,14}. Magavi *et al.* observed cells migrating into the lesioned cortical area. These cells, not present in controls, were apparently recruited from a distance by signals produced as a result of the apoptotic cell death.

Second, we need to identify the signals involved in the neurogenic response. In similar experiments, Macklis and colleagues¹⁵ showed that the expression of several neurotrophic factors (required for neuronal survival, migration and differentiation) is upregulated in interneurons near the degenerating neocortical neurons. Similarly, insults leading to increased neurogenesis in



100 YEARS AGO

The determination of the strength of collateral heredity is a problem of great scientific importance, and it can only be achieved by co-operative action. I have found so many teachers in all classes of schools willing to give disinterested aid in the cause of science that I venture to make a further appeal through *Nature* for more assistance. Besides observations of physical and mental characters, which can be recorded without measurement, my data papers ask for certain head-measurements, which can, following the printed instructions, be taken quite easily. I shall be most glad to send sample papers to any one willing to assist, and if, after considering these, they find themselves able to assist, say by filling in data papers for ten or more pairs of brothers or sisters, I will at once despatch a head-spanner, of which I have several at the present time, free. The head-spanner should not be retained (unless under special circumstances) for more than a few weeks. Where the school is a small one, one master has, as a rule, filled in the papers entirely; in larger schools, one of the science masters, or even the medical officer, has done the head-measurements, and the other data have been provided by house, form or consulting masters. Karl Pearson
From *Nature* 21 June 1900.

50 YEARS AGO

Dictionary of Genetics

This book contains, or attempts to contain, every word connected with genetics in the widest sense. It therefore spreads over (though it does not cover) the whole range of biology; psychology and anatomy, embryology and biochemistry are all represented... For some terms (such as cytomicrosome), the author proceeds by describing *ignotum per ignotius*; others (such as mitoschisis or merostathmokinesis) have never been used except by their anonymous inventors. Let us hope that others again (such as thermocleistogamy, tachyauexesis and spermiocalyptrotheca) never will be used. Others which have a known meaning lose it (such as "mass mutation" in *Enothera*) or do not appear at all (such as "sterility"). Then again, others (such as heterofertilization or parthenogamy) describe rare or even imaginary phenomena. One supreme example which is non-existent as a technical term (perultimate chromomere) appears to be at once misplaced, misspelt and misdefined. From *Nature* 24 June 1950.

the dentate gyrus induce dramatic changes in the expression of many neurotrophic factors¹⁶. The lesion-induced induction of key signalling molecules might represent the re-emergence of a programme that controls cortical development. To use this response to neuronal death in reconstructing functional neuronal circuits, we need to know much more about the sequence of signals that guide the migration, differentiation, integration and connections of neural progenitors.

These results raise the enticing possibility that the brain has a latent capacity for self-repair. However, the neurogenic response observed by Magavi *et al.*¹ was limited, and it is inconceivable that the small fraction of damaged neurons that appeared to be replaced by new neurons would allow significant functional recovery. We do not even know whether the new cells — even though they seemed to form appropriate connections — can take on the function of the cells they replace. There is a long way to go, but learning how to boost and guide neurogenesis from the stem-cell pool might eventually lead to a powerful tool for brain repair in human disorders of the central nervous system. ■

Anders Björklund and Olle Lindvall are at the Wallenberg Neuroscience Center, Lund University, Sölvegatan 17, S-223 62 Lund, Sweden.
e-mails: anders.bjorklund@mphy.lu.se
olle.lindvall@neuro.lu.se

1. Magavi, S. S., Leavitt, B. R. & Macklis, J. D. *Nature* **405**, 951–955 (2000).
2. Gage, F. H. *Science* **287**, 1433–1438 (2000).
3. Gould, E. & Tanapat, P. *Neuroscience* **80**, 427–436 (1997).
4. Bengzon, J. *et al. Proc. Natl Acad. Sci. USA* **94**, 10432–10437 (1997).
5. Parent, J. M. *et al. J. Neurosci.* **17**, 3727–3738 (1997).
6. Liu, J., Solway, K., Messing, R. O. & Sharp, F. R. *J. Neurosci.* **18**, 7768–7778 (1998).

7. Frisén, J., Johansson, C. B., Torok, C., Risling, M. & Lendahl, U. *J. Cell Biol.* **131**, 453–464 (1995).
8. Gensert, J. M. & Goldman, J. E. *Neuron* **19**, 197–203 (1997).
9. Gould, E., Reeves, A. J., Graziano, M. S. A. & Gross, C. G. *Science* **286**, 548–552 (1999).
10. <http://www.sciencemag.org/cgi/content/full/288/5467/771a>
11. Altman, J. *Anat. Rec.* **145**, 573–591 (1963).
12. Kaplan, M. S. *Ann. NY Acad. Sci.* **457**, 173–192 (1985).
13. Weinstein, D. E., Burrola, P. & Kilpatrick, T. J. *Brain Res.* **743**, 11–16 (1996).
14. Nait-Oumesmar, B. *et al. Eur. J. Neurosci.* **11**, 4357–4366 (1999).
15. Wang, Y., Sheen, V. L. & Macklis, J. D. *Exp. Neurol.* **154**, 389–402 (1998).
16. Lindvall, O., Kokaia, Z., Bengzon, J., Elmér, E. & Kokaia, M. *Trends Neurosci.* **17**, 490–496 (1994).

Chemistry

Chirality, magnetism and light

Laurence D. Barron

A tortuous quest involving physicists, chemists and biologists that has endured for over 150 years has finally ended with a paper by Rikken and Raupach¹ on page 932 of this issue. They report the first unequivocal use of a static magnetic field to bias a chemical process in favour of one of

two mirror-image products (left- or right-handed enantiomers). The chemistry of life is homochiral, being based almost exclusively on L-amino acids and D-sugars, and the ability of biological molecules to discriminate between enantiomers is vital for living systems. The importance of handedness in nature is such that scientists have long wondered about its origin, and the process demonstrated by Rikken and Raupach may provide a new clue.

The quest began in 1846 when Faraday made the plane of polarization of a linearly polarized light beam rotate by applying a magnetic field parallel to the beam. This discovery was of fundamental importance because it demonstrated conclusively the intimate connection between electromagnetism and light. But it also became a source of confusion to many scientists who failed to appreciate that there is a distinction between Faraday's magnetic optical rotation and the natural optical rotation discovered three decades earlier by Arago and Biot in certain crystals and fluids. Such natural optical activity is due to the handedness within the microstructure of the crystals and fluids, as Fresnel later showed.

The first to be misled was Pasteur, who in 1848 separated crystals of sodium ammonium tartrate into right- and left-handed forms, which gave equal and opposite natural optical rotations in solution. Following on from this epochal discovery, he attempted to induce handedness in crystals by growing them in a magnetic field², which he mistakenly thought, following Faraday's discovery, to be a source of handedness. But Lord Kelvin, who first introduced the word 'chirality' into science, was under no such misapprehension, and stated quite firmly that "the magnetic rotation has neither left-handed nor right-handed quality, that is to say, no chirality. This was perfectly understood by Faraday, and made clear in

Ecology

Shrimp-eat-shrimp

In 1683, the Ottoman Turks' advance up the Danube was turned back at Vienna. A contemporary crustacean invader, however, is having more success. Jaimie Dick and Dirk Platvoet have discovered that, in the freshwater ecosystems of the Netherlands, the native shrimp *Gammarus duebeni* is being wiped out by a menace from the east, *Dikerogammarus villosus* (*Proc. R. Soc. Lond. B* **267**, 977–983; 2000).

This species is native to eastern Europe and the Ukraine. But *D. villosus* has spread to Western Europe through the Danube–Main canal (which opened in 1992) and appeared in the Netherlands about five years ago. Females, shown here on zebra mussels, are about 15 mm long, males being twice that size.

The alien's method of takeover is nothing if not direct — it eats the natives. In particular, male *D. villosus* consume female *G. duebeni*, which are smaller than males



of the same species and less able to resist attack. The invader is especially destructive because it can feed on its prey between moults, when the exoskeleton is tough, and not just on soft-skinned, recently moulted shrimps.

In some places, *G. duebeni* has already been displaced by a fast-breeding North American species, *G. tigrinus*, which probably reached Europe in ship ballast-water. To an extent it could withstand this assault, because the two species prefer different habitats and salinities which helps to keep them

apart. But *D. villosus* can tolerate a wide range of different conditions, and it is also thought to be responsible for the recent sharp declines in populations of *G. tigrinus*.

As it becomes ever easier for human beings to traverse the globe, exotic animals and plants will be introduced into new environments, both on purpose and as unseen hitchhikers. We face the prospect that ecosystems will become increasingly drab and homogeneous, dominated by a few super-competitive species.

John Whitfield

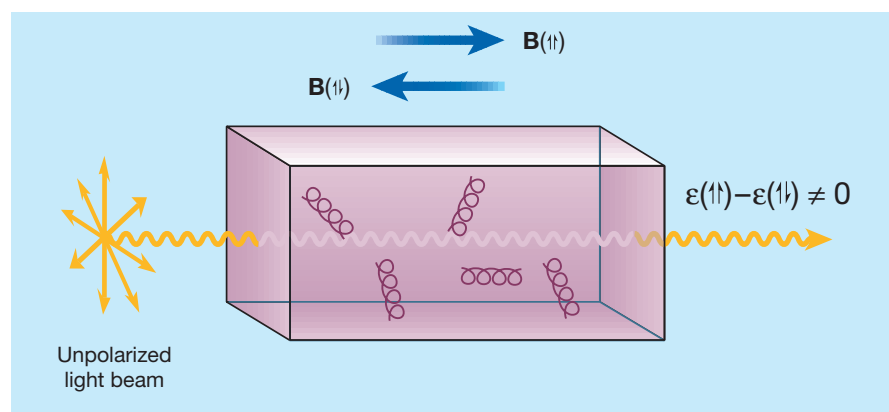


Figure 1 Favouring a lopsided solution. An unpolarized light beam passes through a solution of resolved chiral molecules (represented by small helices) in a static magnetic field either parallel $B(\uparrow)$ or antiparallel $B(\downarrow)$ to the propagation direction. The absorption coefficients $\epsilon(\uparrow)$ and $\epsilon(\downarrow)$ are slightly different owing to magnetochiral dichroism. Rikken and Raupach¹ have now exploited this effect to favour the production of one enantiomer, making it a serious candidate for the source of handedness in nature.

his writings, yet even to the present day we frequently find the chiral rotation and the magnetic rotation classed together in a manner against which Faraday's original description of his discovery of the magnetic polarization contains ample warning.³ Lord Kelvin's admonition was largely ignored, and the next hundred years saw many other futile attempts to use magnetic fields to induce chirality in chemical processes^{4,5}, often motivated, like Pasteur's experiment, to find the source of homochirality in the molecules of life and perhaps even of life's origins.

A new twist to the story appeared in 1982. Wagnière and Meier⁶ predicted that light would be absorbed slightly differently by a solution of chiral molecules if the light beam travelled parallel to an external magnetic field, than if it travelled antiparallel to the field. This small difference in absorption is completely independent of the polarization state of the light beam and so should work with unpolarized light (Fig. 1). This effect, subsequently christened 'magnetochiral dichroism'⁷, depends on a subtle interplay of chiral and magnetic effects on the molecular optical properties, and was observed in 1997 at the Grenoble High Magnetic Field Laboratory by Rikken and Raupach⁸.

Rikken and Raupach¹ have now used magnetochiral dichroism to favour the production of one enantiomer in a photochemical reaction. Their experiment uses the chiral Cr(III)tris-oxalato complex, which is unstable in solution and spontaneously dissociates and re-associates. So at equilibrium there are always equal concentrations of the right- and left-handed enantiomers. This dissociation is accelerated by the absorption of light. The authors show that, in the presence of an unpolarized laser beam travelling parallel to a static magnetic field, a small excess of one enantiomer is produced and maintained, and that, on reversing the magnetic field direction, an equal concentra-

tion of the mirror-image enantiomer results. Their experiment finally achieves Pasteur's aim, albeit in a more subtle fashion than originally conceived by the great scientist.

This work confirms the value of a new definition of chirality that goes beyond Lord Kelvin's original definition (based on mirror reflection) to include time reversal^{4,5,9}, so as to incorporate motion-dependent chirality. This definition provides a rigorous statement of the fundamental symmetry characteristics that external physical fields and forces must have in order to induce absolute enantioselection in all circumstances. Including situations where a chemical reaction has reached thermodynamic equilibrium. According to this new definition, a magnetochiral influence possesses 'true chirality' and

so has the same status as circularly polarized light and the electroweak interaction in its ability to induce absolute enantioselection^{4,5}. These are currently the most favoured explanations for the homochirality of life, and enantioselective photochemistry with circularly polarized light has already been observed experimentally.

On both experimental and theoretical grounds, we now have to seriously consider magnetochiral photochemistry in discussions of the possible origins of biological homochirality¹⁰. This is especially pertinent to fashionable theories suggesting that complex organic molecules could evolve in the ice mantles of dust grains in interstellar space¹¹, because magnetic fields and unpolarized light are more common in the cosmos than circularly polarized light. Furthermore, cosmic magnetic fields lead to partial orientation of the dust grains¹², which may enhance any associated enantioselective chemistry.

Laurence D. Barron is in the Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, UK.

e-mail: laurence@chem.gla.ac.uk

1. Rikken, G. L. J. A. & Raupach, E. *Nature* **405**, 932–935 (2000).
2. Mason, S. F. *Nature* **311**, 19–23 (1984).
3. Lord Kelvin *Baltimore Lectures* (Clay, London, 1904).
4. Avalos, M. et al. *Chem. Rev.* **98**, 2391–2404 (1998).
5. Feringa, B. L. & van Delden, R. A. *Angew. Chem. Int. Edn Engl.* **38**, 3418–3438 (1999).
6. Wagnière, G. & Meier, A. *Chem. Phys. Lett.* **93**, 78–81 (1982).
7. Barron, L. D. & Vrbancich, J. *Mol. Phys.* **51**, 715–730 (1984).
8. Rikken, G. L. J. A. & Raupach, E. *Nature* **390**, 493–494 (1997).
9. Barron, L. D. *J. Am. Chem. Soc.* **108**, 5539–5542 (1986).
10. Wagnière, G. & Meier, A. *Experientia* **39**, 1090–1091 (1983).
11. Cosmovici, C. B., Bowyer, S. & Werthimer, D. (eds) *Astronomical and Biochemical Origins and the Search for Life in the Universe* (Editrice Compositori, Bologna, 1997).
12. Whittet, D. C. B. *Dust in the Galactic Environment* (Institute of Physics Publishing, Bristol, 1992).

Genetics

Reverse gear for *Drosophila*

Barry J. Dickson

These are exciting times for biologists studying the fruitfly *Drosophila melanogaster*. In March came the announcement that the DNA sequence of the fly's genome has been almost completely determined¹. Now, also in *Science*, comes a paper from Rong and Golic² that gives *Drosophilageneticists* the tantalizing prospect of being able to change that DNA sequence almost at will. After nearly a century of classical 'forward' genetics, *Drosophila* now has the necessary gear for 'reverse' genetics as well.

Genetics seeks to bridge the gap between genotype and phenotype. Forward genetics asks which changes in genotype lead to a specific phenotype, while reverse genetics asks how the phenotype responds to specific changes in genotype. Reverse genetics has

long been possible in organisms such as yeast and mice. In these systems, researchers can edit the genome specifically with the aid of linear DNA fragments prepared *in vitro*. These fragments contain the desired sequence changes, flanked by 'homologous' regions, which match the targeted genetic site. Once inside the cell, this exogenous DNA triggers the cellular machinery that normally repairs broken chromosomes or recombines them during meiosis (the formation of eggs and sperm). Directed by the regions of homology, the repair machinery 'recombines' the exogenous DNA into the corresponding chromosomal site. The effect of this sequence change on phenotype can then be assessed.

Why has it been possible to exploit this

pathway in yeast and mice, but not in other organisms? The problem lies not with the repair and recombination machinery, which has been highly conserved during evolution³, but rather in finding an efficient way to get linear DNA fragments into germline cells so that the altered sequence can be passed on to subsequent generations. For unicellular organisms such as yeast, every cell is a germ cell, and it is a relatively simple matter to introduce the exogenous DNA *in vitro*. For the mouse, the problem is solved by genetically modifying embryonic stem cells *in vitro* and injecting them into early embryos, in the hope that these cells will contribute to the germ line. Although this method sparked a revolution in mouse genetics, it has not become universal because of difficulties in obtaining embryonic stem cells from other organisms, even different mouse strains.

Rong and Golic² turn this basic strategy on its head. Rather than trying to introduce linear DNA fragments into germ cells *in vitro*, they have devised a simple method to generate the linear DNA fragments directly in the germ line of the intact organism.

How does it work? The basic idea (Fig. 1) is first to assemble the targeting sequences into a donor element, which is then integrated at a random chromosomal location by standard methods for germline transformation. Then, with the help of a pair of yeast enzymes, also provided transgenically, the linear targeting fragment is everted out of the donor element. This linear fragment can then be recombined into the corresponding chromosomal site.

Rong and Golic tested this system by targeting the *yellow* gene, which controls body colour, in *Drosophila*. They took flies in which the endogenous *yellow* gene was defective as a result of a single base-pair change, and introduced into these flies a donor element containing an intact copy of the *yellow* gene (Fig. 1). When the linear fragment was liberated in the female germ line, it targeted the endogenous locus with astonishingly high frequency: about 1 in 500 progeny carried an intact copy of the *yellow* gene at the endogenous locus. In most cases, the intact copy provided by the donor element had been inserted alongside the defective endogenous copy — precisely what

would be expected following homologous recombination with an exogenous fragment linearized within the region of homology (Fig. 1).

The success of these experiments is certainly cause for celebration. But there is at least one dark cloud on the horizon. Last year, Bellaiche *et al.*⁴ reported a failed attempt to use a similar strategy for gene targeting. Why did Rong and Golic succeed where Bellaiche *et al.* failed? The most likely explanation is that Bellaiche *et al.* attempted targeting in male germ cells after meiosis. In contrast, the targeting events recovered by Rong and Golic most likely occurred before meiosis, and with 16 times higher frequency in females than in males. Both groups targeted genes on the X chromosome, so perhaps the low targeting efficiency in males results from the early inactivation of the X chromosome in the male germ line. Alternatively, this may be a general phenomenon, related perhaps to the lack of meiotic recombination in the male germ line.

There are also other differences between the two studies, including the choice of target locus, the configuration of the donor element, and the extent of homology. It will be important to find out which of these differences really matter.

Although Rong and Golic's method accomplishes the holy grail of gene targeting, it is not yet a general method for reverse genetics. For this, the basic strategy will need to be modified to allow, ideally, any desired sequence change to be introduced at any genetic locus. This will require a general method for selecting targeting events (Rong and Golic suggest a good way to do this). It will also be necessary to distinguish integration at the targeted locus by homologous recombination from non-homologous integration into other genomic locations. Fortunately, there are good reasons to think that homologous recombination will be the more likely fate of a linear fragment in a germ cell^{2,3}.

In recent years, we have come to view forward and reverse genetics as two competing experimental strategies, with different sets of model organisms. But experience with yeast has shown us what great progress can be made once we can explore the terrain between genotype and phenotype using both the unbiased approach of forward genetics and the hypothesis-driven approach of reverse genetics. Now, it seems, this will also be possible in *Drosophila*. One hopes that other organisms will soon follow. ■

Barry J. Dickson is at the Research Institute of Molecular Pathology, Dr Bohr-Gasse 7, A-1030 Vienna, Austria.

e-mail: dickson@nt.imp.univie.ac.at

1. Adams, M. D. *et al.* *Science* **287**, 2185–2195 (2000).
2. Rong, Y. S. & Golic, K. G. *Science* **288**, 2013–2018 (2000).
3. Kanaar, R., Hoeijmakers, J. H. J. & van Gent, D. C. *Trends Cell Biol.* **8**, 483–489 (1998).
4. Bellaiche, Y., Mogila, V. & Perrimon, N. *Genetics* **152**, 1037–1044 (1999).

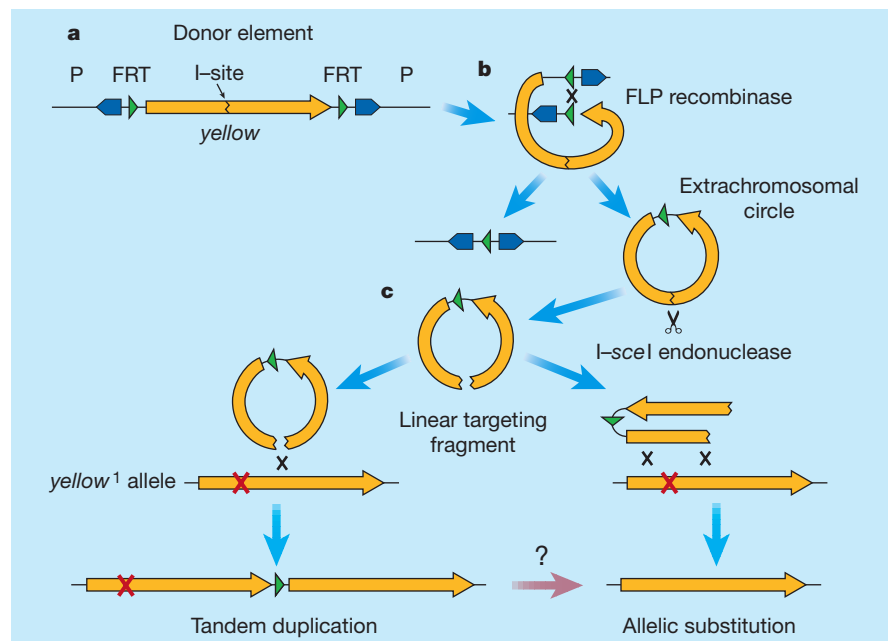


Figure 1 Gene targeting in *Drosophila*. Rong and Golic² altered the sequence of the *yellow* gene in *Drosophila* as follows. **a**, Their donor element contained the intact *yellow* gene, flanked by so-called FRT sites (recognized by the site-specific recombinase FLP) and containing an I-site (recognized by the site-specific DNA-cleaving enzyme I-SceI) in a non-coding region. The donor element and the genes encoding FLP and I-SceI were assembled into transposable elements called P elements to allow them to be inserted into the *Drosophila* genome. The FLP and I-SceI genes were induced by a brief heat pulse during early development. **b**, The FLP protein induced recombination at FRT sites to excise an extrachromosomal circle, which was cleaved by I-SceI at the I-site to generate the linear targeting fragment (**c**). The endogenous locus carried a specific form of the *yellow* gene (*yellow*^b); a mutation of a single base pair (red cross) renders this allele defective and leads to a yellowish body colour. As neither the endogenous locus nor the linearized fragment contained intact copies of the *yellow* gene, homologous recombination between the two was the only way of reconstituting a functional *yellow* gene. Targeting events could thus be recovered by selecting for progeny with normal body colour. Most of these contained tandem duplications; these arise from a single crossover event (black cross). Many allelic substitutions were also recovered. These might result from either double crossover events or the immediate correction of tandem duplications (purple arrow).

Semiconductor physics

Electrons in artificial atoms

Daniel Gammon

With dimensions of only 1 to 100 nanometres and containing somewhere between 10^3 and 10^6 atomic nuclei in a crystalline lattice, semiconductor 'quantum dots' are often described as artificial, solid-state atoms¹. An electron in a quantum dot can be described by a quantum wavefunction that is similar to that used for an electron in a single atom, even though its energy is spread coherently over the lattice of atomic nuclei. The electronic wavefunctions are often labelled with atomic notation (for example, *s* and *p* energy levels), yet the quantum dots are very much solid-state nanostructures, which can be tailored into different shapes. On pages 923 and 926 of this issue, Bayer *et al.*² and Warburton *et al.*³ report optical studies of individual quantum dots and 'quantum rings'. Researchers are fascinated by quantum dots that emit light because such dots are expected to form the basis of a new generation of lasers.

In these studies the semiconductor dots and rings are made from indium arsenide embedded in gallium arsenide. They were grown using techniques developed within the past decade that allow much smaller nanostructures to be created. In these new experiments, electrons are introduced one by one into individual quantum dots while their optical emission is measured with extraordinary precision. Such studies provide new perspectives on the internal quantum-mechanical workings of quantum dots. The ultimate goal is to create useful electronic and optical nanomaterials that have

been quantum-mechanically engineered by tailoring the shape, size, composition and position of various quantum dots.

The importance of semiconductors, large or small, lies primarily in the great change in their properties when the number of active electrons in the material is altered. The electron density can be controlled with high sensitivity through doping, optical excitation or external electric fields. In a semiconductor quantum dot this sensitivity becomes extreme as the electrons are confined in all three dimensions. Complete quantum confinement profoundly affects the electrons, greatly increasing or decreasing the magnitude of their interactions with each other, with their environment and with external fields. Just as in atoms, and in sharp contrast to bulk semiconductor systems, the electrons in the quantum dot exist only in certain quantized energy states (Fig. 1).

It has been instructive to explore the properties of quantum dots as electrons are introduced in a controlled way^{4–9}. For example, adding even a single additional electron takes a significant amount of extra energy because of the repulsion between the negatively charged electrons as they are forced into a smaller volume. One result of this effect, called the Coulomb blockade, is to provide researchers with great control over the number of electrons in a quantum dot. That is, they can easily tune the number by increasing or decreasing the energy they put in.

Currently there is great interest in semi-

conductor quantum dots that interact strongly with light. Optical excitation of a semiconductor leads to the creation of a quasiparticle known as an exciton — a negatively charged electron bound together with a positively charged 'hole'. In contrast to the electrical injection of electrons that leads to the Coulomb blockade effect, a quantum dot remains neutrally charged following optical excitation. Both Bayer *et al.*² and Warburton *et al.*³ study this exciton in detail by measuring the light emitted when the hole and electron recombine. This occurs when extra electrons are electrically injected into the quantum dot³, or when additional excitons are created in a carefully controlled way². In closely related but complementary experiments, the authors build up quantum-dot analogues of the periodic table of atoms as they add electrons or excitons one by one. Repulsive interactions between the particles lead to differences in the amount of energy it takes to create or destroy an exciton.

Quantum confinement of electrons is just one of several ways quantum mechanics reveals itself in semiconductor nanostructures. Another way is through the spin of the electron, which plays a central role in the current studies. An electron with spin behaves like a small magnet, with either 'up' or 'down' spin states. When more than one electron with the same spin exists in the quantum dot, a purely quantum-mechanical interaction, known as an exchange interaction, becomes important. As a result, the energy to create or destroy a second electron (plus its accompanying hole) in the presence of an existing electron also depends on their relative spin.

Electronic states in quantum dots persist for a relatively long time because they interact in a very restricted way with their environment. Normally, such interactions lead to 'decoherence' or destruction of the quantum state. As a result, quantum dots may provide an excellent solid-state system for exploring advanced technologies based on quantum coherence. For example, it may be possible to create and control superimposed or even 'entangled' quantum states using highly coherent laser stimulation¹⁰. External control of the full quantum wavefunction in a semiconductor nanostructure may even lead to revolutionary new applications, such as those involving quantum computing¹¹.

Much more can be done to explore and ultimately harness the quantum-mechanical properties of these artificial solid-state atoms. New research directions are emerging. One that is now in the embryonic stage is the combination of quantum dots into quantum-dot molecules. Many of the quantum-dot systems currently being studied have the potential to be combined into molecular complexes with one-, two- or even three-dimensional structures. One can imagine growing these solid-state atoms or molecules within structures containing

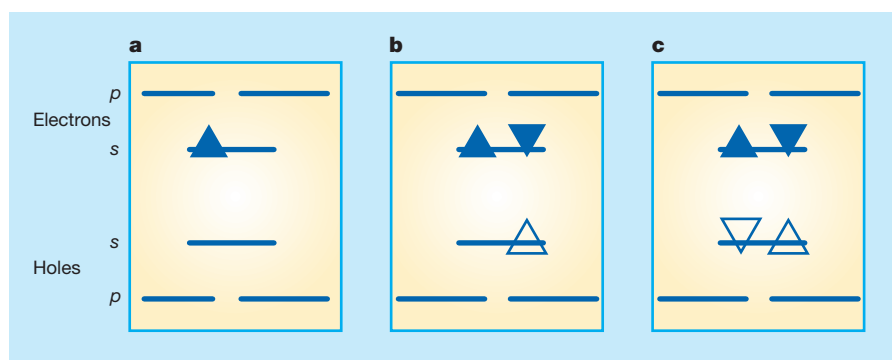


Figure 1 Energy levels in a semiconductor quantum dot. The levels have a shell structure formed of *s* and *p* states, and substructure arising from the spin of the electrons. Electrons (filled triangles) and holes (empty triangles) are shown. a, One or more electrons can be injected into the quantum dots, which fill up the states in a way that minimizes the energy of interactions (Hund's rules of atomic physics). Furthermore, one or more electron-hole pairs (excitons) can be photoexcited. Other possible systems are: b, a single electron-hole pair created in a quantum dot with a single extra electron; c, two electron-hole pairs. In the experiments of Bayer *et al.*² and Warburton *et al.*³, luminescence spectroscopy is used to find the energies (actually the energy differences between states) of these multi-electron quantum dots. In this way, analogues to the periodic table and Hund's rules in quantum dots can be explored.

electronic or magnetic gates and optical cavities, perhaps all connected together by quantum wires.

Quantum dots have great flexibility because their properties can be artificially engineered, but this comes at a price. Nature has given us atoms; scientists must make quantum dots. Further advances in this exciting field of science and technology will depend heavily on the creativity of physicists, chemists and materials scientists who make these tiny structures.

Daniel Gammon is at the Naval Research Laboratory, Code 6876, Building 208, Room 348, 4555 Overlook Avenue SW, Washington DC

20375-5347, USA.

e-mail: gammon@nrl.navy.mil

1. MRS Bull. 23(2) (1998).
2. Bayer, M., Stern, O., Hawrylak, P., Fafard, S. & Forchel, A. *Nature* **405**, 923–926 (2000).
3. Warburton, R. J. *et al. Nature* **405**, 926–929 (2000).
4. Tarucha, S., Austing, D. G., Honda, T., van der Hage, R. J. & Kouwenhoven, L. P. *Phys. Rev. Lett.* **77**, 3613–3616 (1996).
5. Miller, B. T. *et al. Phys. Rev. B* **56**, 6764–6769 (1997).
6. Deckel, E. *et al. Phys. Rev. Lett.* **80**, 4991–4994 (1998).
7. Landin, L. *et al. Science* **280**, 262–264 (1998).
8. Banin, U., Cao, Y., Katz, D. & Millo, O. *Nature* **400**, 542–544 (1999).
9. Findeis, F. *et al. Solid State Commun.* **114**, 227–230 (2000).
10. Bonadeo, N. H. *et al. Science* **282**, 1473–1476 (1998).
11. Berman, G., Doolen, G. D., Mainieri, R. & Tsifrinovich, V. I. *Introduction to Quantum Computers* (World Scientific, Singapore, 1998).

Malaria

A mosquito transformed

Craig J. Coates

Despite the best efforts of hundreds of researchers and many decades of work, malaria remains a scourge that affects millions of people each year. On page 959 of this issue¹, Catteruccia *et al.* describe an approach that should at least speed up understanding of the physiology of the mosquito carriers of the disease and their interaction with the malaria parasite.

Most research to date has centred on the parasite and its development in the human host, but a viable vaccine is yet to materialize.

Furthermore, direct control of mosquito populations is hampered by the spread of insecticide resistance, and attempts to control and treat the disease are commonly foiled by the emergence of drug resistance in the malaria parasites.

Rather less work has been done on the mosquito vectors of malaria and their interactions with the parasites. But Catteruccia *et al.* now report the development of a genetic transformation system for one such vector, *Anopheles stephensi*, one of the major carriers of malaria in urban areas of the Indian subcontinent. In particular, the genetic transformation technology will be used to investigate the molecular aspects of malaria parasite transmission, although it could also lead to new

control mechanisms for the disease. The most ambitious scheme proposed is the production of a genetically engineered strain of mosquito that can no longer transmit the pathogen, and the replacement of the wild population with this 'innocuous' strain^{2,3}.

The process of genetic transformation in insects relies on small circular DNA molecules (plasmids) that are micro-injected into the eggs of the species concerned. Catteruccia *et al.* use a modification of the injection procedure, collecting eggs in a solution containing a chemical that delays egg hardening. This improvement will increase the efficiency of egg hatching and DNA introduction into mosquito species. The plasmids are produced in bacterial cells using recombinant DNA techniques, and they contain mobile DNA sequences called transposable elements, or jumping genes.

The transposable elements can mobilize — jump — from the plasmid DNA molecules and insert themselves into the mosquito chromosomes at random locations. If this event occurs in a nucleus that will form germline cells (sperm or eggs), the resulting offspring will be genetically transformed. The transposable elements can carry additional DNA, in this case a marker gene, so the transgenic individuals can be easily recognized from their non-transgenic siblings.

A transformation system for the fruitfly *Drosophila melanogaster* has been established for almost 20 years and has made possible a vast array of experiments in both insect and human biology. Routine transformation systems for other insects have been developed only much more recently. But in the past two years there has been an exponential increase in the success rate, and ten different non-drosophilid insects have now been transformed.

Although the so-called *P* transposable element is commonly used for *D. melanogaster* transformation, it is inactive in several non-drosophilid insects. So one of the advances in research on such insects has been the availability of several different transposable elements for genetic transformation experiments. Catteruccia *et al.* chose *Minos* from *Drosophila hydei* as their transposable element; *Minos* had already been used to transform the Mediterranean fruitfly *Ceratitis capitata*⁴.

The white-eye gene is typically used as a marker for the transformation of *D. melanogaster*, in which it 'rescues' the eye colour in white-eyed mutant strains. But most non-drosophilid insects do not have any mutant strains. So the marker gene of choice for these species, as used by Catteruccia *et al.*, has been the green fluorescent protein (GFP), which glows green when it is excited with ultraviolet light. A digital image of transgenic mosquito larvae with GFP fluorescence is shown in Fig. 1. With the advances detailed by these authors, transformation

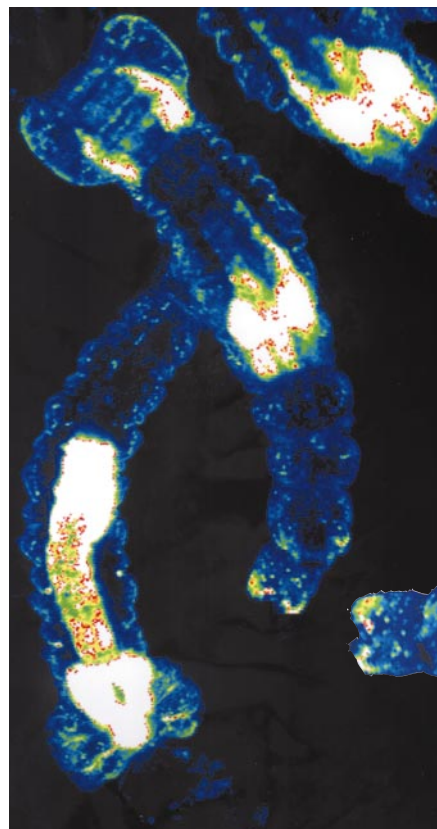


Figure 1 Fluorescence in transgenic mosquito larvae. This is a digital image, taken by Catteruccia and colleagues¹, of larvae expressing the marker gene for green fluorescent protein.

Black areas represent low regions and white areas high regions of green fluorescence. The marker-gene fluorescence enables transgenic mosquitoes to be distinguished from their non-transgenic siblings.

Oceanography

Given a twirl

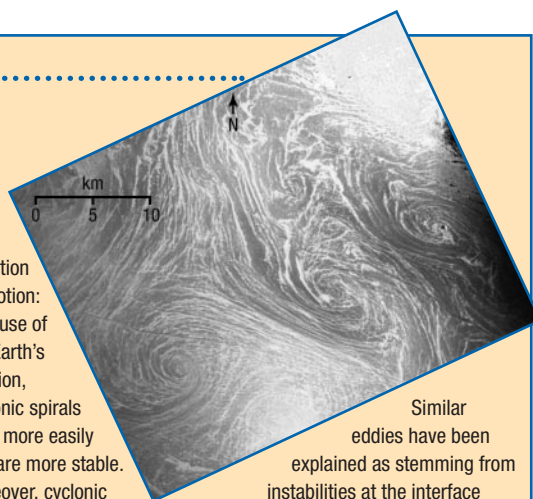
Only with the advent of space oceanography have spiral structures in the sea become recognized phenomena. They have proved rewarding, if difficult, subjects for research.

Walter Munk and colleagues have analysed 400 photographs from space that show these attractive patterns on the ocean surface (*Proc. R. Soc. Lond. A* **456**, 1217–1280; 2000). They restricted their investigation to small spirals, 10–25 km in diameter, that last for only about a day, excluding features such as the rings in the Gulf Stream that can be up to 300 km across and last for years.

It turns out that most spirals rotate in the same direction as storm cyclones: anticlockwise and clockwise in the Northern and Southern Hemispheres, respectively. Munk *et al.* find that all potential mechanisms favour the same

direction of motion: because of the Earth's rotation, cyclonic spirals form more easily and are more stable. Moreover, cyclonic rotation is more likely to be visible, because streaks on the sea surface that can be twisted into spirals are more likely to occur where cyclones are created.

For much the same reasons, the Mediterranean in autumn is particularly rich in visible spirals: the seasonal, strong winds, followed by calm conditions, are ideal generators of streaks. The picture here shows an example of fully developed spirals off Crete in October 1984.



Similar eddies have been explained as stemming from instabilities at the interface between two water masses of different densities. But that cannot apply to many of the images studied by Munk *et al.* Rather, they think that the driving force is often turbulence generated at the boundary of two water masses moving relative to each other.

As the authors remark, however, there is certainly more than one way to make an ocean spiral. Which may explain why they are such common features of the seascape. **Heike Langenberg**

technology should soon be applied to the main vector of human malaria, *Anopheles gambiae*. This is the principal vector of malaria in sub-Saharan Africa, where more than 90% of all malaria cases — and an estimated one million deaths each year — occur.

Not all forms of a mosquito vector species can transmit malaria parasites; those that cannot are called refractory strains. Advances in understanding the molecular basis of the mechanisms involved will provide molecules that can be introduced into a receptive strain to make it refractory. Furthermore, advances in understanding insect immune systems are providing potential target genes that can be transformed into a mosquito strain to prevent the development and transmission of malaria parasites. Researchers can also borrow molecules of immunity, such as antibodies, from vertebrate systems — through a variety of molecular methods, the binding domains of antibodies that recognize malaria parasites can be cloned as small pieces of DNA and then incorporated into the mosquito genome.

Although it is likely that a refractory mosquito strain will be created before too long, the question of how it would be introduced into the wild remains. It is unlikely that enough mosquitoes could be mass-reared to allow a simple replacement of the wild population, even assuming that the engineered

strain had an equal chance of surviving under natural conditions. But perhaps the ability of transposable elements to spread through natural populations could be used as a driving force to spread the refractory strain⁵. Research into cytoplasmic incompatibility phenomena, such as that caused by the symbiotic bacterium *Wolbachia*, may provide another possible driving force⁶. Cytoplasmic incompatibility causes sterility under certain mating conditions, aiding the transmission and spread of the bacteria.

The consequences and risks associated with the field release of a transgenic mosquito will require careful assessment before the approach could become a reality. Whatever the final outcome, however, the successful transformation of a mosquito vector of human malaria is a notable advance in our ability to combat this devastating disease. ■

Craig J. Coates is at the Center for Advanced Invertebrate Molecular Sciences, Department of Entomology, Texas A&M University, College Station, Texas 77843-2475, USA.
e-mail: c-coates@tamu.edu

1. Catteruccia, F. *et al.* *Nature* **405**, 959–962 (2000).
2. Collins, F. H. & James, A. A. *Sci. Med.* **3**, 52–61 (1996).
3. O'Brochta, D. A. & Atkinson, P. W. *Sci. Am.* **279**, 90–95 (1998).
4. Loukeris, T. G., Livadaras, I., Arca, B., Zabalou, S. & Savakis, C. *Science* **270**, 2002–2005 (1995).
5. Kidwell, M. G. & Ribeiro, J. M. C. *Parasitol. Today* **8**, 325–329 (1992).
6. Curtis, C. F. & Sinkins, S. P. *Parasitology* **116**, S111–S115 (1998).

Daedalus

Solid explosion

Fire-retardant paint is a cunning product. When strongly heated, it swells up into a frothy char which seals gaps and insulates its substrate from the heat. Daedalus is now taking the idea to extremes. He is devising a polymeric composition loaded with separate tiny particles of a mild explosive, of the type used in automotive air-bags. When detonated, it will expand suddenly and violently into a solid resembling foamed polystyrene.

The chemistry of this novel product is quite challenging. Each particle of explosive must expand into a gas bubble, transiently melting and deforming the polymeric solid around it. The heated viscous thixotropic polymer must be tenacious enough not to fragment under this sudden shock, although it will be hot and soft at the moment of explosion, and will inflate in a pneumatic manner. It will incorporate an added monomer which rapidly sets in the heat, absorbing heat by its reaction and reinforcing the final expanded solid. The mild explosive must be sensitive enough for one particle to set off others nearby, but not enough to make the composition too dangerous to handle. It will be tricky to develop, but the final product, DREADCO's 'Solid Charge', will have many uses.

For a start, it will wonderfully extend the air-bag form of impact protection. Painted on the sharp edges of a vehicle inside and out, and set off by cunning sensors, it could cushion impacts to passengers inside and pedestrians outside. In block form, Solid Charge could offer other forms of safety. While most explosives blow things open, it could blow things shut. It could block pipes, tunnels or entrances against leaks or spreading fire. It will pose no fire or smoke hazards itself, as it seals in its own fumes — useful in mining as well as domestic protection. It might even defeat burglars, by sealing their way in or out, or suddenly filling an opened safe or closet with a protective polymeric mass encapsulating the contents.

Exploded inside in a tough folded plastic envelope of adequate volume, Solid Charge will extend air-bag technology in another way. It will expand to take the shape of the inflated envelope. It will suddenly create an 'instantaneous object' of defined shape. Daedalus likes the idea of a self-inflating instantaneous rubber dinghy for marine emergencies. Once inflated it could not spring a leak, and would ride solidly, without the queasy flexibility of conventional rubber craft. **David Jones**

Obituary

Kent R. Wilson (1937–2000)

Kent Wilson, who died on 27 March, developed innovative laser techniques to probe the molecular dynamics of chemical and biochemical reactions. A remarkably wide-ranging physical chemist, he also contributed much seminal work, both theoretical and experimental, to several other fields.

In the past few years, Wilson's group at the University of California, San Diego, opened up three frontiers. Each required development of powerful new apparatus, exemplifying Wilson's architectural vision. One machine achieves 'quantum control' of photochemical processes, using iterative feedback to automatically tailor laser excitation pulses for optimum effect. Another, the culmination of long endeavour, enables direct observation of atomic motions by means of ultrafast X-ray diffraction and spectroscopy. Most recently, the group has built a laser microscope that can obtain very-high-resolution images of living tissue without killing the cells.

Wilson was born on 14 January 1937 in Philadelphia, and grew up in a nearby village, Bryn Gweled Homesteads. This was a mostly Quaker community, which his parents had helped to found, designed to foster communal living. Encouraged by scientists in the village, by the age of eight Wilson had made a rudimentary radio receiver. Later, his family spent winters in Washington DC, where his father headed a Quaker organization working for disarmament and civil rights. Wilson became an avid visitor to the Smithsonian Museums and the Library of Congress; he credited that with "saving my academic career from ruin" because he "did very little of the assigned homework in school".

At Harvard College, Wilson began as a government major studying chiefly history and political science. However, he was captivated by a chemistry course given by the fabled Leonard Nash. That led Wilson to switch his major to chemistry and physics. After completing a bachelor of arts degree in 1958, he went to the University of Strasbourg for a year and wrote a *Diplôme* thesis in economics. He then decided to pursue graduate study in chemistry, and moved to the University of California at Berkeley. Susceptible to evangelical fervour, he joined my fledgling research group just as we were beginning crossed-molecular-beam studies of the reactions between alkali atoms. His doctoral thesis, completed in 1964,



Inspiring architect of laser chemistry

described the design and construction of a state-of-the-art apparatus and presented data mapping out the generic relations between product angular distributions and electronic structure of the reactant molecules.

In 1965, Wilson moved to San Diego and began building his own research group. He undertook the development of laser experiments to elucidate the dynamics of molecular photodissociation. The techniques of ordinary spectroscopy put almost exclusive emphasis on stable excited states, but in molecules there are many more dissociative excited states. Wilson's technique of photofragment spectroscopy opened up a vast lode of information about such states and repulsive intermolecular forces.

Although his later research explored quite diverse areas, his *leitmotif* remained molecular dynamics. In the 1980s, by combining ultrafast laser experiments with computer simulations, Wilson markedly advanced the understanding of reaction dynamics in solution, bringing out connections to gas-phase dynamics. In the 1990s, his ultrafast X-ray laser apparatus revealed unexpected dynamical aspects of the transition between solid and liquid phases, in a phenomenon termed 'non-thermal melting'. In striving for his ultimate goal of quantum control, he hoped to conduct 'conversations' with molecules through laser light pulses, enabling the molecules to teach experimenters how to negotiate a desired

chemical choreography of atomic motions.

For Wilson, teaching and mentoring were adventures in both intellectual and social dynamics. He dubbed a unique part of his research group the 'Senses Bureau', to which he recruited first-year undergraduates "young enough so that they do not yet know the meaning of the word impossible". Over 35 years, they created software and hardware for computer animation and produced many educational films and interactive virtual-reality devices. A classic film explicated protein synthesis with the aid of 200 dancers, poetry and a jazz-rock band. The bureau also had a large part in Wilson's studies of air pollution in California and Mexico, and in the use of chemical archaeology to trace trade routes in Africa.

In mid-career, Wilson became increasingly frustrated with the time and tribulation of applying for research grants. On a sabbatical leave in 1985, he came up with an extraordinary solution, discovering "a mapping between parts of physics and economics". Boldly betting on his theory, over the next few years he invested in the stock markets of 25 countries. This succeeded so well that he was able to resign all his federal grants, thereafter using his own funds to support his sizeable research group and provide laser and computer instrumentation.

Late in 1998, Wilson learned he had incurable, late-stage prostate cancer. For the remaining 14 months of his life he dealt with this as his "most challenging experiment", organizing a team of biochemists to work with physicians in pursuing new therapeutic approaches. Last December, the *Journal of Physical Chemistry* published a special issue honouring Wilson. He contributed an autobiographical essay, a poignant and wise "love letter to posterity". The closing paragraph describes his joy in visiting wild flowers along a favourite trail, akin to his pleasure in the blooming of his scientific projects and progeny. In this, and in other ways, Wilson was an heir of Max Planck, whose study of the interaction of light and matter led to the discovery of quantum physics a century ago. In a letter to a friend, Planck wrote:

*What you have picked, what I have picked
These we will bind together,
Entwining thus a fair bouquet
From gifts we send each other.*

Dudley Herschbach

Dudley Herschbach is in the Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138, USA.
e-mail: herschbach@chemistry.harvard.edu

Whale songs lengthen in response to sonar

Male humpbacks modify their sexual displays when exposed to man-made noise.

There is growing concern about the effects of man-made noise on marine life. In particular, marine mammals that use sound to communicate, navigate, and detect predators and prey may try to avoid loud sound sources up to tens of kilometres away¹. Here, in a study conducted in cooperation with the US Navy², we show that the singing behaviour of male humpback whales was altered when they were exposed to LFA (low-frequency active) sonar. As the song of these whales is associated with reproduction³, widespread alteration of their singing behaviour might affect demographic parameters, or it could represent a strategy to compensate for interference from the sonar.

During the breeding season male humpback whales sing long, complex songs that are thought to be sexual displays³. Songs consist of a series of themes, progressing in

a predictable order, that may repeat for several hours⁴. We used a small observation vessel to find singing humpbacks and conduct focal sampling⁵, recording behaviour before, during and after playback. (Strictly speaking, we have evaluated the additional impact that LFA sounds have on a singing whale that is already being followed.)

We recorded the vocal behaviour of each focal singer continuously for several hours using a towed, calibrated hydrophone array⁶. When the whale was at the surface, observers sampled visible behaviour. Photographs of fluke and dorsal fin features confirmed the whale's identity throughout each follow⁷. At least two songs were recorded before the observation vessel requested the US Navy R/V *Cory Chouest* to transmit ten (in one case four) 42-s LFA signals at 6-min intervals. The sonar was broadcast at less than full strength, and no focal singer was exposed to a signal louder than 150 dB (with respect to 1 μ Pa).

Sixteen singers were followed during 18 playbacks. In nine follows, the whale sang continuously throughout the playback; in four the singer stopped when he joined other whales (typical of normal social interaction); and in five the singer stopped, presumably in response to the playback. We recorded at least one complete song in all conditions from six individuals, and pooled the songs of each of the two individuals that were subjects in two experiments. For these six whales, we measured the duration and theme structure of song spectrograms, comparing song duration in the three conditions using analysis of variance⁸.

On average, humpback whales' songs were 29% longer during LFA playbacks

(Fig. 1) — a particularly strong result, given the low power of the test and small sample size⁹. Song duration returned to normal after exposure, suggesting that this response has a limited duration. There was little difference in the likelihood of an aberrant theme transition across exposure conditions ($\chi^2 = 3.273$, $P = 0.195$), indicating that long songs resulted from longer themes within a normal song structure. Across the six singers, maximum received level of the sonar at the whale did not correlate positively with either the increase in mean song duration from pre-exposure to exposure condition ($r = -0.90$) or with the subsequent decrease from exposure to post-exposure condition ($r = -0.63$).

We suggest that humpbacks sang longer songs during LFA sonar transmissions to compensate for acoustic interference. Our study shows that it is possible to measure the behavioural responses of individual whales in controlled experiments at sea.

Patrick J. O. Miller*, **Nicoletta Biassoni***, **Amy Samuels*†**, **Peter L. Tyack***

*Biology Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

†Daniel F. and Ada L. Rice Conservation Biology and Research Center, Chicago Zoological Society, Brookfield, Illinois 60513, USA

1. Richardson, W. J., Greene, C. R. Jr, Malm, C. I. & Thomson, D. H. *Marine Mammals and Noise* (Academic, San Diego, 1995).
2. Waters, M. A. *US Federal Register* **61**, 37452–37453 (1996).
3. Tyack, P. L. *Behav. Ecol. Sociobiol.* **8**, 105–116 (1981).
4. Payne, R. & McVay, S. *Science* **173**, 585–597 (1971).
5. Altmann, J. *Behaviour* **49**, 227–267 (1974).
6. Miller, P. J. & Tyack, P. L. *Deep-Sea Res. II* **45**, 1389–1405 (1998).
7. Katona, S. *et al.* in *The Behavior of Marine Mammals* (eds Winn, H. E. & Olla, B. L.) 33–44 (Plenum, New York, 1979).
8. Zar, J. H. *Biostatistical Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1984).
9. Lehmann, E. L. *Ann. Math. Stat.* **29**, 1167–1176 (1958).

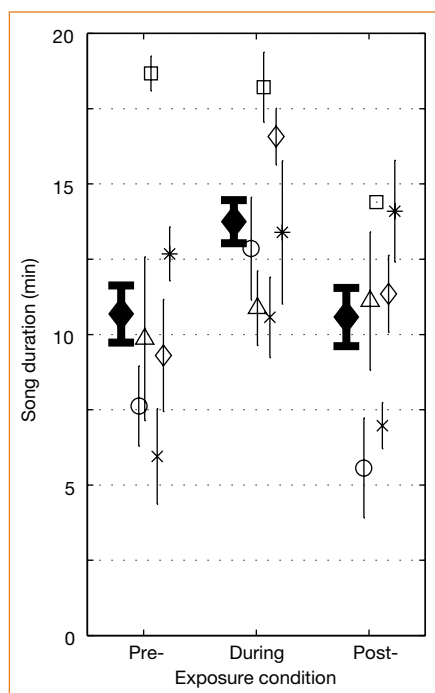


Figure 1 Duration of songs ($\pm$ s.e.m.) produced by humpbacks before, during and after exposure to LFA sonar transmissions (bold, filled diamonds, mean of all six singers; other symbols, individual singers). The maximum received level of the sonar at the whale ranged from 130 to 150 dB re 1 μ Pa. Songs were grouped in the exposure condition if a sonar transmission occurred at any point during the song. The average number of songs per singer in the pre-exposure, exposure and post-exposure conditions was 3.2, 4.7 and 3.8, respectively. Differences were assessed using a mixed-model analysis of variance treating exposure condition as a fixed factor, whale identity as a random factor, and each song duration as an independent observation. The effect of exposure condition on song duration was statistically significant at $P = 0.047$ ($F_{2,10} = 4.200$, power = 0.50, $n = 6$).

Nutrition

Antioxidant activity of fresh apples

Vitamin C is used as a dietary supplement because of its antioxidant activity, although a high dose (500 mg) may act as a pro-oxidant in the body^{1,2}. Here we show that 100 g of fresh apples has an antioxidant activity equivalent to 1,500 mg of vitamin C, and that whole-apple extracts inhibit the growth of colon- and liver-cancer cells *in vitro* in a dose-dependent manner. Our results indicate that natural antioxidants from fresh fruit could be more effective than a dietary supplement.

Apples of the Red Delicious variety were

extracted using 80% acetone and their content of phenolics and flavonoids determined^{3,4}: the extracts contained 290.2 ± 4.2 and 219.8 ± 1.8 mg phenolics, and 142.7 ± 3.7 and 97.6 ± 3.9 mg flavonoids per 100 g apples with and without skin, respectively. There are known to be more phenolics in the skins of apples than in the flesh, and quercetin glycosides are found only in the skins⁵.

We measured the total antioxidant activity of apples by using the total oxyradical-scavenging capacity (TOSC) assay⁶ and found that apples with skin had a higher TOSC value than apples without skin (Fig. 1a). The total antioxidant activity of 1 g apples with skin was 83.3 ± 8.9 TOSC (μ mol vitamin C equivalents) — that is, the

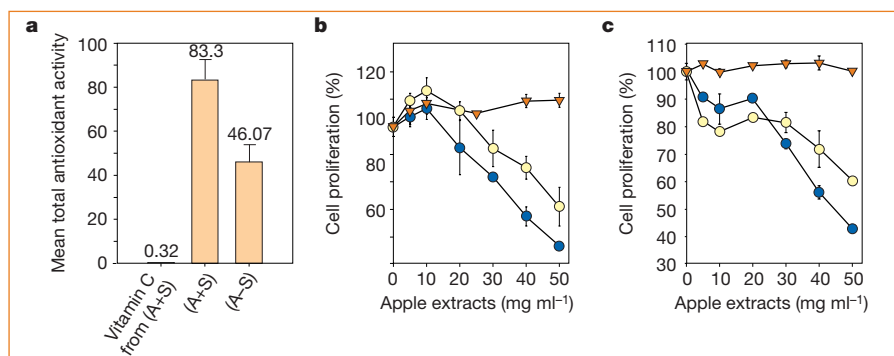


Figure 1 Antioxidant activity of apples and their effect on tumour-cell proliferation *in vitro*. **a**, Mean total antioxidant activity expressed by total oxyradical-scavenging capacity (TOSC; μmol vitamin C equivalents per g) assay for 1 g apple with skin (A + S), and for apple without skin (A – S). **b**, Inhibition of proliferation of Caco-2 colon-tumour cells by extracts of apple with and without skin. **c**, Inhibition of proliferation of HepG2 liver-tumour cells. Control samples were assayed as apple extracts, but they contained only vitamin C at the same concentration that exists in apples (0.057 mg g^{-1}). Cell proliferation was determined by using the MTS assay¹⁰. Blue circles, apple with skin; yellow circles, apple without skin; orange triangles, control extracts.

antioxidant value of 100 g apples is equivalent to 1,500 mg of vitamin C. Given that the average vitamin C content in fresh apples with skin is 5.7 mg per 100 g (ref. 7) and that the total antioxidant activity of 0.057 mg vitamin C (in 1 g of whole apples) is only 0.32 TOSC (Fig. 1a), then almost all of the antioxidant activity in apples must be due to phytochemicals.

We treated a colon-cancer cell line, Caco-2, with extracts equivalent to 0, 5, 10, 20, 30, 40 and 50 mg ml^{-1} apples for 96 hours (the treatment time for maximal response). Cell proliferation was inhibited in a dose-dependent manner after exposure to apple-extract concentrations above 20 mg ml^{-1} (Fig. 1b): at 50 mg ml^{-1} , inhibition was $43 \pm 1\%$ and $29 \pm 4.1\%$ for apples with skin and for apples without skin, respectively.

We also tested the effect of apple extracts on the proliferation of another cancer-cell line, HepG2 human liver-tumour cells. We found that apple extracts at 50 mg ml^{-1} inhibited the proliferation of these cells as well, by $57 \pm 0.21\%$ and $40 \pm 0.64\%$ for apples with and without skin, respectively (Fig. 1c). The extracts of apple with skin could thus significantly (*t*-test, $P < 0.031$) reduce tumour-cell proliferation compared with extracts of apples without skin. No cytotoxicity of the apple extracts was seen at any of the concentrations tested (data not shown).

We suggest that this strong inhibition of tumour-cell proliferation *in vitro* could be due to apples' combination of phytochemicals (phenolic acids and flavonoids), as these are natural antioxidants. It has been proposed that the consumption of whole fruits may provide the antioxidant balance needed to quench reactive oxygen species⁸ which have been implicated in tumorigenesis⁹. Phytochemicals in apples other than ascorbic acid seem significantly to enhance their antioxidant properties and their capacity to inhibit the proliferation of tumour cells *in vitro*.

Marian V. Eberhardt, Chang Yong Lee, Rui Hai Liu

Department of Food Science, 108 Stocking Hall, Cornell University, Ithaca, New York 14853-7201, USA

e-mail: RL23@cornell.edu

1. Podmore, I. D. *et al.* *Nature* **392**, 559 (1998).
2. Herbert, V. *et al.* *J. Nutr.* **126**, 1213S–1220S (1996).
3. Singleton, V. L. & Rossi, J. A. Jr *Am. J. Enol. Viticult.* **16**, 144–158 (1965).
4. Zhisen, J., Mengcheng, T. & Jianming, W. *Food Chem.* **64**, 555–559 (1999).
5. Burda, S., Oleszek, W. & Lee, C. Y. *J. Agric. Food Chem.* **38**, 945–948 (1990).
6. Winston, G. W., Regoli, F., Dugas, A. J., Fong, J. H. & Blanchard, K. A. *Free Rad. Biol. Med.* **24**, 480–493 (1998).
7. USDA Nutrient Database for Standard Reference (1998).
8. Wang, H., Cao, G. & Prior, R. L. *J. Agric. Food Chem.* **44**, 701–705 (1998).
9. Lunec, J. *J. Int. Fed. Clin. Chem.* **4**, 58–63 (1992).
10. Cory, A. H., Owen, T. C., Bartrop, J. A. & Cory, J. G. *Cancer Commun.* **3**, 207–212 (1991).

Botany

Constraints to growth of boreal forests

Understanding how the growth of trees at high latitudes in boreal forest is controlled is important for projections of global carbon sequestration and timber production in relation to climate change. Is stem growth of boreal forest trees constrained by the length of the growing season when stem cambial cells divide¹, or by the length of the period when resources can be captured²? In both cases, the timing of the thaw in the spring is critical: neither cambial cell division nor uptake of nutrients and carbon dioxide can occur while the soil is frozen. Here we argue, on the basis of long-term observations made in northern Saskatchewan and Sweden, that the time between the spring thaw and the autumn freeze determines the amount of annual tree growth, mainly through temperature effects on carbon-dioxide uptake in spring

and on nutrient availability and uptake during summer, rather than on cambial cell division.

From tree-ring analysis across several sites in the Siberian subarctic, Vaganov *et al.*¹ provide evidence that annual variability in mean ring width is determined by the date of the thaw through its influence on the date of cambial initiation, as well as by temperature during the subsequent early growing season. Variability in annual net ecosystem production (NEP) is also largely determined by the timing of the thaw³, which enables the NEP to switch immediately and rapidly from a daily loss to a daily gain of CO_2 .

Our own observations on black spruce trees at the Boreas Southern Study Area⁴ show that when air temperature exceeds -1°C and the overlying snow starts to melt, meltwater percolates down into the soil, the temperature of the upper soil horizons rises towards 0°C , and the switch from a small net daily loss of carbon to a large net gain occurs over just a few days. In boreal conifers, the availability of soil water is a prerequisite for the recovery of photosynthetic capacity in spring and early summer².

The effect of frozen soils on annual CO_2 uptake by Norway spruce at 64°N is particularly dramatic because, before the thaw, daily solar radiation is already substantial and effectively being wasted from the perspective of CO_2 capture⁵. Thus, CO_2 uptake is synchronized and strongly stimulated by the thaw, and afterwards, once a critical temperature sum is reached, cambial activity and NEP increase together as the temperature rises.

Such observations related to the poor growth of trees in boreal forests have led to the presumption that their growth is constrained by temperature. By contrast with temperate and tropical forests⁶, boreal forest trees are small in relation to their age and coniferous boreal forests have a very low net primary production of about 2.5 tonnes of carbon $\text{ha}^{-1} \text{yr}^{-1}$ (refs 7,8). To investigate the extent to which low temperature is the primary controlling variable, a long-term nutrient-optimization and irrigation experiment on Norway spruce was set up at Flakaliden (64°N) in Sweden⁹. Since 1987, we have applied complete fertilizer daily through every growing season either in irrigation water or as a single solid dose at the start of the growing season. We found that growth on the fertilized plots (with or without irrigation) increased by a factor of four¹⁰ (Fig. 1a), so air temperature cannot be the major direct constraint on tree growth.

However, temperature may be influencing tree growth indirectly through the length of the growing season and by its effects on decomposition of soil organic

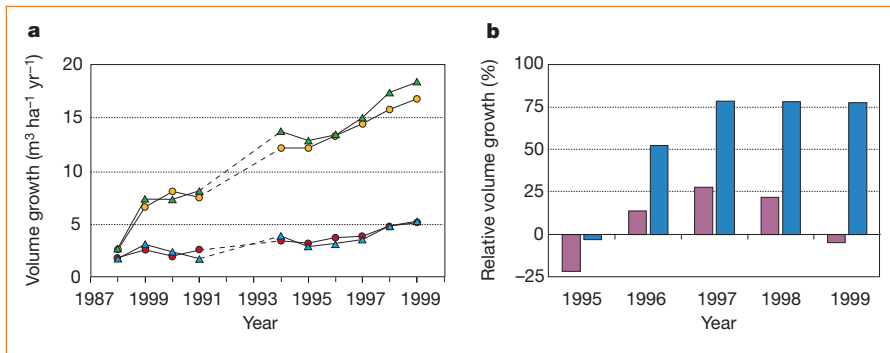


Figure 1 Annual increment of stem volume in stands of young Norway spruce during a nutrient-optimization experiment, and relative effect of soil warming on the annual volume production at Flakaliden in northern Sweden. **a**, In the nutrient experiment, the treatments were: control (red circles), irrigation (blue triangles), solid fertilization (yellow circles) and combined irrigation and fertilization (green triangles); all started in 1987^{9,10}. **b**, In the soil-warming experiment, volume growth values are given in per cent of that in 1994, before the start of the experiment. The low production in 1995 was the effect of massive cone production. Purple bars, non-heated control plots; blue bars, heated plots.

matter and mineralization of soil nutrients. A long-term soil-warming experiment (an increase of 5 °C at 10 cm depth, through the growing season²) at Flakaliden is revealing after five years that concentrations of foliar nutrients are increased and that there is a significant (more than 50%) increase in stem-wood growth of the trees on the heated plots (Fig. 1b). Thus, year-to-year variability in temperature could be influencing stem-wood growth through the availability of nutrients.

Although this experiment makes this point well, a note of caution is appropriate. After five years of warming, CO₂ efflux from the forest floor in warmed and non-warmed plots differs by less than 10%, which is not significant. The evidence from our experiment thus contradicts the idea that the projected rise in temperature is likely to lead to forests that are now carbon sinks becoming carbon sources in the foreseeable future. There is likely to be a limit to the amount of soil organic matter that is readily metabolizable in response to an increase in soil temperature, as well as an acclimation taking place in the temperature sensitivity of both autotrophic and heterotrophic respiration.

We conclude that the effects of temperature on the growth of trees in the boreal forest are mediated by the capture of CO₂ and nutrients much more strongly than by other physiological processes. Consequently, the variability from year to year in the net ecosystem and net primary production and in stem wood growth should be interpreted, in the first instance, in terms of the ability of the forest to capture carbon and nutrients.

Paul Jarvis*, Sune Linder†

*Institute of Ecology and Resource Management, University of Edinburgh, The King's Buildings, Mayfield Road, Edinburgh EH9 3JU, UK
†Department for Production Ecology, Swedish University of Agricultural Sciences, PO Box 7042, SE-750 07 Uppsala, Sweden

1. Vaganov, E. A., Hughes, M. K., Kiryanov, A. V., Schweingruber, F. H. & Silkin, P. P. *Nature* **400**, 149–151 (1999).
2. Bergh, J. & Linder, S. *Glob. Change Biol.* **5**, 245–253 (1999).
3. Goulden, M. L. *et al. Science* **279**, 214–217 (1998).
4. Jarvis, P. G. *et al.* 1996 Boreas data, Oak Ridge National Laboratory archive (Oak Ridge, TN).
5. Bergh, J., McMurtrie, R. E. & Linder, S. *Forest Ecol. Mgmt* **110**, 125–139 (1998).
6. Malhi, Y., Baldocchi, D. D. & Jarvis, P. G. *Plant Cell Envir.* **22**, 715–740 (1999).
7. Gower, S. T. *et al. J. Geophys. Res.* **102**, 29029–29041 (1997).
8. Jarvis, P. G. *et al. J. Geophys. Res.* **102**, 28953–28966 (1997).
9. Linder, S. *Ecol. Bull. (Copenhagen)* **44**, 178–190 (1995).
10. Bergh, J., Linder, S., Lundmark, T. & Elfving, B. *Forest Ecol. Mgmt* **119**, 51–62 (1999).

Vaganov and Hughes reply — The model we used, which is driven only by meteorological variables, reproduces the decadal timescale variations in relative tree-ring width indices that have been observed at three taiga sites in the Siberian subarctic separated by 85 degrees of longitude¹. In our study, absolute differences in mean growth rate between sites, related to slowly varying site conditions such as nutrient status, are removed by converting to dimensionless indices with a mean of unity. The results from the fertilization experiment on young plantation spruce trees described by Jarvis and Linder are thus not in disagreement with our major findings.

Jarvis and Linder suggest that the effect of soil temperature on the timing of cambial initiation is indirect, rather than direct. We acknowledge that it is not possible to say whether the effects of temperature and moisture are direct or indirect using the methods we adopted, especially for wood increment, which results from a combination of different processes at the whole-tree and tissue levels. We would point out, however, that our simple model does a remarkable job of accounting for the changes observed over much of this century (see Fig. 4 of ref. 1), and that there is little room for improvement on this.

Their comments are relevant to the study of changes that the global carbon

cycle might undergo in a changing climate, although we would counsel caution in assuming that any single mode of response will apply to the wide diversity of conditions found in the boreal forest. The problem is how to include the proposed indirect effects in quantitative models of the effect of climate variability on the variability of tree growth. In particular, what is the characteristic time of the response of wood formation to various environmental changes?

We know that, in the case of the meteorological variables used in the model we adopted, characteristic times generally range between a few days and one or two growth seasons. Thus, in some extremely cold years the same tree at the same average level of nutrients in the soil forms a tree ring an order of magnitude narrower than in warm years. Fertilization increases stem-volume growth and stand productivity². This effect is revealed clearly, with a delay of some years, in non-permafrost soils. All the sites we investigated were on deep permafrost with a thin active layer.

Under these circumstances, then, can nutrient supply play a quantifiable role in the variation of interannual and inter-decadal growth rate in response to climate variability and change, even in the same site and in the same tree? In a warm year, the active layer will be deeper and warmer, leading to a higher rate of organic decomposition and additional turnover of nutrients for growth³. Whereas the direct effect of meteorological variables on wood formation is close to immediate, there is likely to be a significant delay in the operation of indirect effects related to nutrients, because the active layer is deepest and warmest in the later part of the season, namely in late July and August^{4,5}. Thus, cumulative decomposition is greatest when wood formation is already almost complete, delaying effects at least until the next year. These delays will probably make it difficult to disentangle the direct and indirect effects of meteorological factors on tree radial increment. In the meantime, our model accounts for most of the observed inter-decadal variability in relative tree-ring growth.

Eugene A. Vaganov*, Malcolm K. Hughes†

*Institute of Forest, Siberian Branch, Russian Academy of Science, Akademgorodok, 660036 Krasnoyarsk, Russia

†Laboratory of Tree-Ring Research, University of Arizona, Tucson, Arizona 85721, USA

1. Vaganov, E. A., Hughes, M. K., Kiryanov, A. V., Schweingruber, F. H. & Silkin, P. P. *Nature* **400**, 149–151 (1999).
2. Buziykin, A. I. (ed.) *The Forests of the Middle Angara Region* (Nauka, Novosibirsk, 1977; in Russian).
3. Goulden, M. L. *et al. Science* **279**, 214–217 (1998).
4. Abaimov, A. P. *et al. Lesovdenie* (Russ. J. Forest.) (in the press; in Russian).
5. Abaimov, A. P. *et al. Proc. Sixth Symp. on Joint Siberian Permafrost Studies between Japan and Russia in 1998, Sapporo, Japan* 18–19 (Hokkaido Univ., Sapporo, 1999).

The genetic legacy of the Quaternary ice ages

Godfrey Hewitt

School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK

Global climate has fluctuated greatly during the past three million years, leading to the recent major ice ages. An inescapable consequence for most living organisms is great changes in their distribution, which are expressed differently in boreal, temperate and tropical zones. Such range changes can be expected to have genetic consequences, and the advent of DNA technology provides most suitable markers to examine these. Several good data sets are now available, which provide tests of expectations, insights into species colonization and unexpected genetic subdivision and mixture of species. The genetic structure of human populations may be viewed in the same context. The present genetic structure of populations, species and communities has been mainly formed by Quaternary ice ages, and genetic, fossil and physical data combined can greatly help our understanding of how organisms were so affected.

The study of palaeoclimates is a particularly active research field that is producing much data and increasingly coherent explanations. The Earth's climate became cooler through the Tertiary (65 million years (Myr)) with frequent oscillations that increased in amplitude and led to the series of major ice ages of the Quaternary (2.4 Myr to the present). The evidence for such global fluctuations in climate comes particularly from cores of the sea bed, lake bottoms and ice sheets, which are analysed for carbon and oxygen isotopes, radiolarian species, pollen types and other biological and physical signatures^{1,2}.

Ice ages and species distributions

While the Antarctic ice cap grew from the Oligocene (35 Myr), the Arctic ice cap became established about 2.4 Myr ago, the beginning of the Quaternary. From then until 0.9 Myr ago, the ice sheets advanced and receded with a roughly 41,000-yr (41-kyr) cycle; thereafter they have followed a 100-kyr cycle and become increasingly dramatic. Such periodicity suggests a controlling mechanism, and the Croll–Milankovitch theory proposes that the regular variations in the Earth's orbit around the Sun are the pacemakers of the ice-age cycles^{1,2}. The main orbital eccentricity has a 100-kyr cycle, variation in the Earth's axial tilt has a 41-kyr cycle, and precession due to the Earth's axial wobble has a 19–23-kyr cycle; these all modify the insolation of the Earth and the energy it receives. Much energy is transported by the oceanic circulation system, and the interaction of orbital variation and currents leads to significant climate changes^{2,3}.

It is now possible to extract ice cores of ~2 km in length and analyse the annually layered snow for entrapped gases, isotopes, acidity, dust and pollen. Some recent cores sample ice over 400 kyr (ref. 4), but most go back ~125 kyr from the present to the previous (Eemian) interglacial. Long pollen cores stretching back to 400 kyr are also becoming available⁵. Along with other measures, these are providing a detailed picture of the last glacial cycle and glimpses of the preceding ones (Fig. 1). The Greenland (Arctic) and Vostok (Antarctic) ice cores are particularly informative, offering fine temporal resolution and continuity². This has revealed surprising oscillations of climate on a millennial scale within the main 100-kyr cycle. The Greenland Ice Core Project (GRIP) identifies some 24 interstadials through the last ice age with average temperature rising rapidly by ~7 °C over just decades. Further ice and sediment cores from around the world are demonstrating the global scale of these major climatic events^{6,7}. As more long cores of ice, sediment and pollen become available, it will be possible to produce a synthesis of the effects of these rapid climatic switches on plant and animal life worldwide⁸.

These severe climatic oscillations produced great changes in species distributions, and these have been described in some detail, particularly from the fossil records of pollen and beetles in Europe and North America^{1,9}. Species went extinct over large parts of their range, some dispersed to new locations, some survived in refugia and then expanded again, and this must have occurred repeatedly.

During major glaciations the polar ice sheets spread considerably, and temperature, marine and vegetation zones were compressed towards the Equator². Mountain blocks like the Alps, Andes, Rockies and Yakutsk ranges also had considerable glaciation, so that the large volume of accumulated ice reduced sea levels by about 120 m (ref. 10). This produced land bridges in several parts of the world (Fig. 1). In warmer parts species descended from mountains. Tropical rainforest was restricted and dissected, and there was extension of deserts and savannah¹¹. It is apparent that these major climatic shifts were felt differently across the globe owing to regional differences in landform, ocean currents and latitude. Furthermore, species responded individually, and their range changes were particular to local geography and climate.

Genetic consequences

In Europe and North America, an extensive network of pollen cores tells us that species now inhabiting boreal and temperate regions had their ice-age refugia south of the ice and permafrost. Post-glacial expansion into new territory was previously suggested to be important in the geographic distribution of population and species genomes¹². It was remarkably rapid for many species, and the suddenness of the large climatic shifts recorded in the recent ice cores helps explain this. Populations at the northern limits of the refugial range would have expanded into often large areas of suitable territory. This leading edge expansion would probably be by long-distance dispersers that set up colonies and rapidly expanded to fill the area before others arrived. This would be repeated many times over a long colonizing route, and these founding events would lead to loss of alleles and homozygosity¹³. Modelling and simulations of such range expansion show that leptokurtic dispersal produces large areas of homozygosity as compared with normal or stepping-stone modes, and these homogeneous areas persist in space and increase with time¹⁴. Secondary oscillations within the main expansion increase the effect. Rapid colonization by this leading edge model in any part of the world should produce areas with reduced genomic variability.

Where the postglacially colonized regions are known from the fossil record (in Europe and North America), a growing number of studies show them to have lower genetic diversity¹⁵, and a fine

analysis of 41 North American fish species demonstrates this clearly¹⁶. These studies also reveal that equivalent genomes occupy larger ranges in these colonized areas, as produced in simulation modelling¹⁴.

An interesting corollary of this leading edge expansion is that when populations have filled the space it is much more difficult for those behind them to advance, because they must disperse and reproduce logistically, and not exponentially like the original scattered colonists¹³. This high-density barrier will mean that boundaries between expanded edge and blocked interior genomes will tend to persist for some time and be seen in present surveys^{15,16}. It is also likely that during a range change a retreating rear edge will suffer shrinkage, dissection and extinction, so that the last surviving population should be severely bottlenecked. This could be in the vanguard of the recolonization.

With expansion of range there will also be selection and adaptation to different environments and new neighbours^{13,15}. For example, the reproductive biology of many plants is suited to different light and temperature regimes across their present range from south to north¹⁷, which has been produced with postglacial colonization. The same is true of insects, and particularly nice demonstrations are seen in recent invasions with accompanying modifications of life history, such as that of the USA by the cornborer moth *Ostrinia nubilalis*¹⁸. Theoretical and experimental work in flies has shown an increase in genetic variance after severe bottlenecks, probably due to the reassortment of epistatic interactions¹⁹. In the pitcher-plant mosquito, *Wyeomyia smithii*, which ranges from Florida to Labrador, a reduction in allozyme diversity and heterozygosity in the postglacial expansion range in the north is reflected in an increase in additive genetic variance for development time and photoperiod response²⁰. The genetic architecture has been remodelled by the dynamics of colonization. Such reorganization during colonization would compound divergence between population genomes. With repeated climatic oscillations and range changes, a population may pass through many such adaptations and reorganizations and its

genome structure could diverge considerably.

The rapid expansion across large parts of Europe and North America was not followed by all species, and was probably much slower in some other regions, particularly the tropics. This would maintain a larger effective population size and retain much more genetic diversity. Various parts of the same species range may have been colonized at different rates owing to physical barriers or prior inhabitants, producing distinct genetic structures. Mountain blocks in southern temperate regions and the tropics would allow slower altitudinal shifts in range, tending to retain allelic diversity. Their varied topography also tends to subdivide the species into populations that may evolve independently with only occasional gene flow, perhaps only every glaciation or interstadial. It would seem that the interpretation of present genetic structure needs to consider such interaction of biology, geography and climatic shifts.

Suitable DNA markers

Modern DNA technology allows genetic diversity to be measured as single-base changes in many individuals, and a variety of sequences have been used (Table 1). As the Quaternary is 2.4 Myr old, most DNA sequences will diverge little over the ice ages, and few new mutations will distinguish postglacial haplotypes. Consequently, differences in such markers between postglacial population genotypes are likely to be due to the sorting of mutations that have mostly originated in the more distant past. Some regions in the D-loop of human mitochondrial (mt) DNA have been shown to be hypervariable, providing useful distinguishing mutations. Some care is called for in the interpretation of patterns^{21,22}, and there has been considerable development of techniques for probing population history. Population bottlenecks are inferred by mismatch comparisons²³, aspects of demography and phylogeography can be discerned using nested clade analysis²¹, and spanning haplotype networks can reveal both demographic and geographic history²⁴. These are now being applied effectively to human populations.

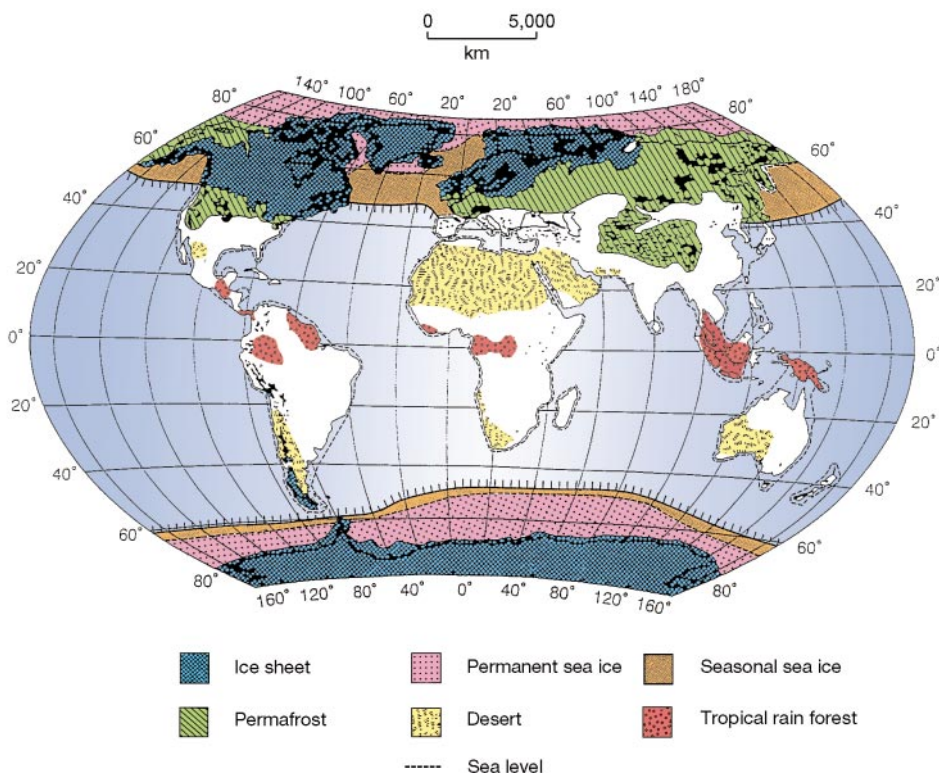


Figure 1 The maximum extent of ice and permafrost at the end of the last ice age 20,000 yr BP. The lowered sea level, large deserts and main blocks of tropical forest are indicated. Modified with permission from ref. 2.

European case studies

Research accumulated before the discovery of polymerase chain reaction (PCR) showed that the distribution of many species are subdivided by narrow hybrid zones; these were produced by the meeting of two diverged genomes as they expanded their ranges from separate glacial refugia¹². With methods such as PCR, it is now possible using DNA similarity to show from which refugia particular genomes emerged to cover their present distribution. There are now a few species that have been studied using suitably variable and discriminating sequences in a large enough number of samples from across their range^{15,25}. Wide-range coverage is important to produce a full phylogeography and because the effects of leading and rear edge, refugia and altitudinal shifts are expected to be different. Twelve such cases have been collated²² in which likely postglacial colonization routes from their refugia across Europe can be deduced. Although each species colonization has its own colonization history and geographic pattern of genomes, three broad patterns are evident—the grasshopper, the hedgehog and the bear—which may serve as paradigms.

Chorthippus parallelus, the common meadow grasshopper of Europe was analysed using a unique non-coding nuclear sequence²⁶, which revealed that its species genome was divided into at least five main geographic regions. Across all northern Europe the haplotypes showed little diversity and were similar to those in the Balkans, clearly indicating a postglacial expansion from a Balkan refugium. Turkey, Greece, Italy and Spain each contain a large portion of unique haplotypes, which shows that they contained refugia and that these populations were limited to these southern regions. Hybrid zones have been described between the Spanish and French genomes and the Italian and French/Austrian genomes along the Pyrenees and Alps where the expanding genomes met²⁷, and may also occur where other genomes meet in southern Europe. Pollen analysis shows that the vegetation that currently supports these grasshoppers was restricted to parts of southern Europe during the last glacial maximum, and notably demonstrated rapid postglacial advance in the east (Fig. 2).

The alder, beech and crested newt have genome patterns that are broadly similar to that of the grasshopper, with their deduced colonizations dominated by a Balkan expansion²². The rapid colonization of northern Europe by alder from a north Balkan source seen in the pollen fits well with the genetic evidence, but the chloroplast DNA haplotypes show several more distinct southern refugia indiscernible in pollen²⁸.

Erinaceus europeus, the western European hedgehog is parapatric with its sibling species *Erinaceus concolor* in the east. After an intriguing allozyme study, its phylogeography has been examined with sequences from the cytochrome (cyt) b mtDNA gene²⁹.

Parsimony and distance trees both show considerable divergence between the two species, and surprisingly reveal distinct clades within these. The depth of divergence between these lineages suggests separations several million years ago, possibly at the beginning of the Quaternary. The hedgehog genome across Europe is thus divided into three principal north/south strips, with a fourth clade in Turkey and Israel. This pattern is quite different from the grasshopper and points to the colonization of northern Europe from three glacial refugia, a western/Spanish one, a central/Italian one and an eastern/Balkan one (Fig. 2). Turkey was also a refugium, and more data may show further subdivisions, particularly in the south and east. Colonization patterns similar to the hedgehog are seen in the oaks and silver fir, with genomes from several southern refugia contributing to the postglacial advance to the north. The populations of oaks emerging from their last glacial refugia probably contained different mixtures of more ancient chloroplast DNAs²². A narrow hybrid zone in Finland between western and eastern cytotypes indicates colonization by very different routes³⁰.

Ursus arctos, the brown bear, has had its broad European distribution reduced by man to a few isolates in the west, with larger populations in Scandinavia, eastern Europe and Russia. Consequently it is the object of conservation efforts. Analysis of the mtDNA control region revealed distinct eastern and western lineages³¹ and *U. arctos* appears to have colonized most of Europe from an Iberian and a Caucasian/Carpathian refuge²² (Fig 2). These expansions met in central Sweden, probably as the last of the Scandinavian ice cap melted ~9 kyr ago, and formed a hybrid zone. The western mtDNA lineage is subdivided, but the clades from Italy and the Balkans did not expand into northern Europe, possibly prevented by the prior occupation of a rapid expansion from the other refugia. From present genetic data, it is not clear exactly from where the eastern expansion emanated, nor how far west it spread across Europe. This latter issue may be addressed through studies of ancient DNA. The bear pattern with its east/west embrace of Europe is similar to that found in shrews and water voles^{22,25}. The presence of several small mammal hybrid zones in the same place in Sweden as that of the bear³² clearly indicates distinct eastern and western colonizations as the ice melted. All three species had distinct genomes that remained in Italy and probably other parts of southern Europe. Studies on other species are in progress, and it will be informative to compare them with these broad paradigm patterns. We can expect individual variants, and maybe a strikingly different one exists.

Europe—emerging features

It appears that the Balkans were a source for all species in the east and for many species in the west; the grasshopper/alder/beech/newt pattern would seem to be common; conversely, Italian genomes rarely populated northern Europe, for example, hedgehog and oak. It seems that the ice-capped Alps were an initial barrier to their northward expansion, while the Pyrenees apparently acted as a barrier in fewer cases, for example, grasshopper, alder and beech. No firm generalizations can yet be made about colonization pattern from species attributes, such as dispersal powers, generation time or habitat adaptation. More species need to be examined.

When these expanding genomes met they formed hybrid zones, and many of these have been described across Europe¹². These appear to cluster in the Alps and central Europe, and to some extent in the north Balkans and Pyrenees^{13,15,22,25}. Remington³³ called such clusters “suture zones” from his review of North American species. The neat cluster of zones in central Sweden was probably produced by the final melting of the Scandinavian ice cap, which allowed eastern and western colonists to meet (Fig. 3). For the Balkans, more genetic data is needed; they appear genomically diverse and dissected, and the source of most northward postglacial expansions.

Britain received oaks, shrews, hedgehogs and bears from Spain, but grasshoppers, alder, beech and newts from the Balkans. The

Table 1 The rates of evolution of various DNA markers

DNA marker	Organism	Average rate of divergence*
Nuclear DNA		
Synonymous (silent) sites	Mammals	0.7
	<i>Drosophila</i>	3.12
	Plant (monocot)	0.114
Intron	Mammals	0.7
Chloroplast DNA		
Synonymous (silent) sites	Plant (Angiosperm)	0.024–0.116
Mitochondrial DNA		
Protein-coding regions	Mammals	2.0
	<i>Drosophila</i>	2.0
	Shrimps	1.4
	Human	14
	Human	17.5
COI	Human	270
D-loop	Human	270
Synonymous (silent) sites	Plant (Angiosperm)	0.01–0.042
Microsatellite		
	Human	5.6×10^{-4}
	<i>Drosophila</i>	6.3×10^{-4}

* Rates of evolution of various DNA markers are given as per cent divergence per million years, which for microsatellites is in units of mutation per locus per gamete per generation (for details, see refs 84–86).

mixture in Sweden is even more complex. This causes one to pause in considering coevolutionary relationships and community structure, both in terms of their evolution and stability. Furthermore, there is fossil evidence for very different communities in refugia and during the many range changes.

Partly because of these glacially induced range changes, northern Europe has less genetic variety than southern Europe in terms of numbers of species, subspecific divisions and allelic diversity: “southern richness to northern purity,” so to speak¹⁵. Rapid expansion may well explain the loss of alleles in northern populations, whereas the varied topography of southern parts provided suitable habitats through several glacial cycles, thereby allowing the divergence and accumulation of several genomes^{15,22}. This process may continue to produce species ultimately. It has significant conservation implications.

Some species have managed to maintain themselves in southern Europe for many ice ages, whereas others have colonized or recolonized more recently (Table 2). The toad *Bombina* and hedgehog may well have been diverging for 5 Myr, but the eastern and western genomes of the grasshopper and the bear probably first entered their refugia about 0.4 Myr—four ice ages ago—a time of Red Sea salinity¹⁰ and a long cold period. It may be that a number of species became extinct in Europe, even in their southern peninsular refugia. They recolonized these after this and have been expanding, contracting and diverging since. Genomes surviving in these refugial areas thus diverge by repeated allopatry, protected by hybrid zones^{12,13}, and may speciate in one or over many ice ages.

Other parts of the world

The type of consideration using molecular phylogeographies for Europe can be applied to other parts of the world, bearing in mind the differences in geography and climatic history, to seek possible generalities. In such a review only some points can be highlighted.

North America. Although a continent of similar size and at similar latitude, North America differs significantly from Europe in geographical features that should have affected the structuring of genomes and species by the climatic oscillations of the Quaternary. Altitudinal shifts, north/south range changes and recolonization from outside would be different. The ice sheets advanced further south to 40°N, and the tundra belt was much narrower in the east. The mountainous west was more varied with ice, tundra, pluvial lakes and deserts, and the Bering land bridge connected Alaska and Siberia².

There is an increasing flow of phylogeographic papers using molecular markers in North American species³⁴, and only a few cases can be mentioned here. The seminal comparative studies were on marine and coastal species in southeastern USA³⁵, where northern and southern DNA lineages made contact around Florida.

Interestingly, this well studied region contains one of the “suture zones between recently joined biotas” that Remington³³ proposed for terrestrial organisms, and parapatric mtDNA clades have been described for several land species here³⁶. These terrestrial zones may be analogous to those in southern Europe where the northern expanding taxa blocked the advance of those from the south. A similar explanation is possible for the suture zones that run from New England to the Dakotas, which broadly coincide with the edge of the ice sheets. The southeastern states also show a suture zone for freshwater fishes, where western and eastern drainages and genomes meet around the Alabama–Georgia line in the southern Appalachians³⁶. The turtles of this region provide a fine example of how phylogeographic concordance can demonstrate shared biogeographical histories^{37,38}. The considerable species richness and substructure in southeastern USA was probably generated by the survival and divergence of genomes in refugia, as ranges changed through repeated glacials and interglacials, with changes in sea level, ocean currents, peninsular climate and water drainages. Such southern richness is found in Europe, but the Appalachians run north/south and possibly played a different role from the European mountains.

Many North American studies report lower genetic diversity in northern populations that have expanded from refugia south of the ice sheets^{15,39,40}, and also greater clade area in fish¹⁶. As the ice melted it produced large flooding proglacial lakes, which could have dispersed aquatic organisms extensively in both Nearctic and Palaearctic. Plant and animal species in the Pacific northwest show the southern richness and northern purity produced by the retreat of the ice sheets^{39,40}, and greatly strengthen the conclusions drawn from the northern colonization of Europe^{14,15}. Interestingly, several widespread North American birds have low genetic diversity across their range reflecting a rapid postglacial expansion, but several other species show genome subdivision in the west and southwest USA⁴¹. Indeed, these parts are now and were during glaciations dissected by mountains and deserts, which produced complex habitat shifts, long-term isolation and the genome divergence that is seen in a number of species⁴².

The Bering land bridge was free of ice during the glaciation, and genetic phylogenies show that Berigian refugial genomes colonized deglaciated regions of North America in some fish and beetles^{16,43}. Phylogeographic information is becoming available for a range of other Arctic organisms, some circumpolar, which shows that their genomes have been fragmented by glaciations to produce subspecies and in some cases species. These include guillemots⁴⁴, dunlins⁴⁵, reindeer⁴⁶ and bears^{47,48}.

It would seem useful to see North America as principal regions differing in geography, climate and hence species history and genomic pattern. The deglaciated northern parts with a distinctive northwest region most resemble Europe, whereas the well-studied southeast would seem richer than southern Europe because of the

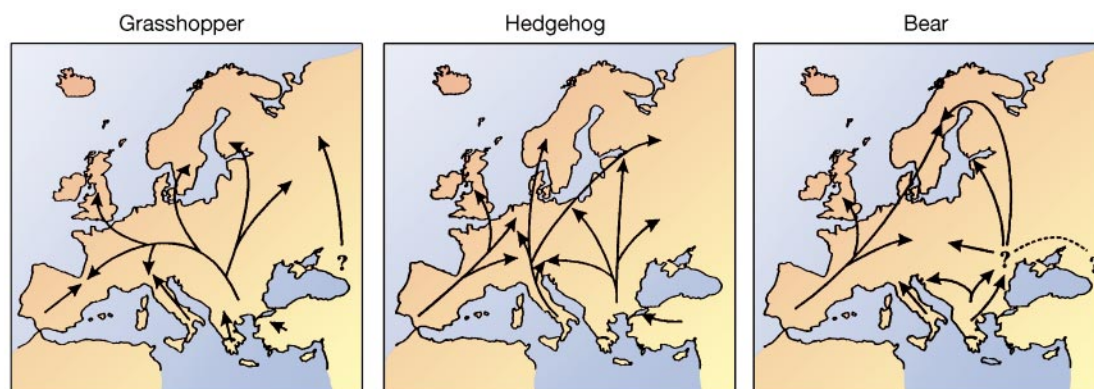


Figure 2 Three paradigm postglacial colonizations from southern Europe deduced from DNA differences for the grasshopper, *Chorthippus parallelus*, the hedgehog, *Erinaceus*

europeus/concolor, and the bear, *Ursus arctos*. The main refugial areas, Iberia, Italy, the Balkans and Caucasus, contributed differently to the repopulation of northern parts.

presence of the Mediterranean and North Africa at similar latitudes. The complex southwest has no real European analogue, except perhaps Iberia. More full species phylogeographies are needed to build regional syntheses and indicate their general principles.

Tropical latitudes. Conditions during the last ice age were colder and drier in the tropics, extending deserts and savannah while reducing rain forests. Unfortunately, these have a poor pollen record, but some recent studies in Amazonia and Queensland suggest that the rain forest species survived in many local wet places and along moist gullies^{2,49}. Apparently, temperatures were 6°C lower in the mountains, and forest species descended to the lowlands. The tropics are particularly speciose, and so the combined investigation of palaeoclimatology, palaeontology and molecular phylogeography, as in Europe and North America, promises to be interesting and important.

There are already a few revealing studies in Australia, Amazonia, Africa and southeast Asia. Climatic oscillations have expanded and dissected the strip of rain forest in northeast Queensland and with it the species there. Several birds, reptiles and frogs show low mtDNA diversity as the result of population contractions, whereas some have retained overall diversity through different haplotypes surviving in several patches. Their phylogeographies reveal a concordant divide between north and south clusters, which fits with an initial separation at the start of the Quaternary⁵⁰. Amazonia poses a particularly difficult task owing to its large size and inadequate palynology, taxonomy and sampling⁵¹. So the comparative phylogeography of 35 species of small mammals for 15 genera using mtDNA is most welcome⁵². Surprisingly, many taxa have deep sequence divergence dating to the Pliocene. The Pleistocene Andes uplift producing the Iquitos Arch across the Jura river is coincident with phylogenetic breaks in 11 of the taxa, although genome divergence at the major rivers themselves does not support the river barrier hypothesis for Amazonian speciation.

There are studies emerging for a range of organisms in Central America and the Caribbean, particularly in birds and fish. For freshwater fishes, the picture looks complex with several overlaid cycles of colonization and lineage extinction from the late Miocene through the Pleistocene⁵³.

Savannah. While the tropical forests were reduced during the Quaternary ice ages, the drier climate allowed grassland and savannah to expand, so that species associated with these habitats

had a different cycle of contraction into refugia. Several African bovids inhabit the Savannah, and recent mtDNA phylogeographies of the hartebeest, topi and wildebeest show how the climatic oscillations affected their distribution and genomic divergence⁵⁴. The first fossils of these species are reported from the early Quaternary, and distinct monophyletic mtDNA clades indicate that their Africa-wide distribution has been reduced several times to a few refugia. The asymmetric gene tree topology of the wildebeest and a decline in sequence diversity from South Africa to east Africa clearly suggest that it colonized from South Africa in recent cold dry periods. The first fossils are from east Africa, but this population probably did not survive and the region was recolonized. Clearly, the careful combination of climatic, fossil and genetic evidence can provide a much clearer picture of the evolution in these habitats also.

Tropical mountains. Like the small mammals in Amazonia, bird species in lowland tropical forest of South America and Africa are quite old with more than 6 Myr of DNA divergence^{55,56}. In contrast, mountain regions in the tropics contain clusters of recently diverged lineages along with older species. So it has been proposed that these mountains provide through the ice ages a stable moist habitat, in which older species may survive and new lineages be generated. More evidence for tropical mountains as centres for bird speciation through the Pleistocene climatic changes comes from mtDNA divergence in greenbuls from east Africa⁵⁷ and spinetails from the Andes⁵⁸. As suggested for southern Europe, these refugial and speciation properties of the mountains may derive from their topographic variety, which allows habitats and lineages to persist by altitudinal shifts and also diverge because of distributional dissection.

Wallacea. The lowering of the sea level during the ice ages produced a number of land bridges and sea barriers. The Sunda shelf in southeast Asia is perhaps the largest of these, and linked much of Malaysia and Indonesia west of the Wallace line to the Asian mainland. The straits of Malacca between Malaya and Sumatra are both narrow and shallow, so even small falls in sea level would produce a land bridge. Consequently, this region will have been dissected and rejoined on many occasions by climatic oscillations back into the Pliocene affecting the distribution and evolution of species.

Some mtDNA investigations show deep divergences that would suggest Pliocene and early Pleistocene radiations. In the case of the spineless hedgehog, *Hylomys*, there are distinct mainland and island clades⁵⁹, whereas the river catfish ancestral lineages form three regional groups⁶⁰. The mtDNA of the striped rabbit, *Nesolagus*, in Sumatra is very different from that of newly discovered representatives in Laos⁶¹. These data suggest there has been no admixture of

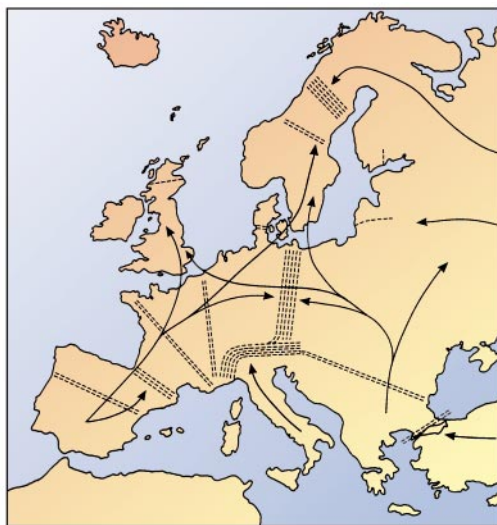


Figure 3 The general position of some well-known hybrid zones in Europe, which show major clustering in Scandinavia, central Europe and the Alps. Other clusters are apparent in the Pyrenees and the Balkans. These suture zones are caused by commonalities of ice-age refugia, rate of postglacial expansion and physical barriers. There is further subdivision in the southern regions.

Table 2 DNA sequence divergence and maximum time of separation in species groups that colonized Europe after the last ice age

Organism	DNA sequence	Divergence	Maximum age	Refugia*
<i>Bombina orientalis</i>	mt RFLP	9.4%	5 Myr	(I) B B
'Fire bellied toad'				
<i>Erinaceus europaeus</i>	mt cyt b	6–12%	3–6 Myr	S I B
'Hedgehog'				
<i>Triturus cristatus</i>	mt RFLP	4–8%	2–4 Myr	S I B
'Crested newts'				
<i>Arvicola terrestris</i>	mt cyt b	4–7.6%	2–4 Myr	S I B
'Water vole'				
<i>Crocidura suaveolens</i>	mt cyt b	3–6.4%	1.5–3.2 Myr	S (I) B
'White toothed shrew'				
<i>Mus musculus</i>	mt RFLP	3.4%	1.7 Myr	W & E
'House mouse'				
<i>Microtus agrestis</i>	mt RFLP	2%	1 Myr	(S) B E
'Field vole'				
<i>Sorex araneus</i>	mt cyt b	1–3.8%	0.5–2 Myr	S (I) B
'Red toothed shrew'				
<i>Ursus arctos</i>	mt control region	2.7–7%	0.35–0.85 Myr	S (I) B
'Brown bear'				
<i>Chorthippus parallelus</i>	mt 6.7 kb	0.7–0.9%	0.3–0.5 Myr	(S) (I) B
'Meadow grasshopper'				

* Deduced southern refugia of distinct genomes are given. S, Iberia; I, Italy; B, Balkans; E, east; W, west. Brackets show ones that did not expand out (see ref. 22 for details). RFLP, restriction-fragment length polymorphism.

populations during the many Pleistocene land-joining. Other diverged species have been discovered from the mountains of this region.

A different pattern is seen in the mtDNA of the flying fruit bat, *Eonycteris*, along the Indonesian archipelago. Sequence divergence suggests events in the Pleistocene and also reflects the sea crossing distance between islands, with a decreasing diversity from ancestral western populations to colonized eastern ones. There are indications that this may be another emerging concordant pattern in the fauna of these islands⁶².

Seas and oceans. For several marine species the Indo/west Pacific region, which includes the Sunda Shelf and Indonesia to the Torres Straits, marks a phylogenetic disjunction between Indian and Pacific lineages and taxa. This includes butterflyfish, starfish, damselfish and coconut crab, and coastal mangroves and fishes^{63–66}. The depth of mtDNA and allozyme divergence would place the effective separation of these sister taxa in the Pleistocene, which is when the sea level was lowered during ice ages and the Asian–Australian archipelagos were more of a barrier. As with terrestrial and freshwater organisms, the Indo/west Pacific seas are rich in species, which may be due to the many changes in distribution caused by the sea level oscillations in the topographically complex archipelagoes. It may function as a species source and refuge in a similar way to mountainous regions in lower latitudes.

From studies so far it seems that species in the large oceans show little geographic genetic differentiation, as exemplified in the butterfly fish in the tropical Pacific⁶³. In temperate waters, coastal sardine taxa from separate parts of north and south Pacific and Indian Oceans from mtDNA sequences would appear to have diverged within the past 0.5 Myr, with each geographic form carrying the genetic signature of rapid expansion from a founding colony⁶⁷. A survey has revealed that pelagic fishes in general, with their high dispersal and large but fluctuating population sizes, show low genetic differentiation among geographic regions; the 21 species ranged from tuna to anchovies⁶⁸. Such low genetic diversity might be expected for temperate species that have undergone range changes due to the ice ages, but it is not yet clear how this compares with their refugial areas at lower latitudes and in the Tropics.

Quaternary speciation

Increasing climatic oscillations during the past 3 Myr have driven many range changes in all parts of the world. It was necessary to move, adapt or go extinct, and present lineages had the ability and luck to survive such shifts. Each time they would colonize new territory, face new environments and meet new neighbours. These challenges would cause genomes to diverge, both through selection and chance, and ultimately speciate¹³. The time taken to speciation and hence the rate of speciation are currently of some interest.

It was generally thought that the Pleistocene increased the initiation of species, but a compilation of molecular data for divergence in subspecies and species complexes showed that species were formed through the Pliocene and Pleistocene¹⁵. A more recent analysis of available molecular differences among lineages and sister species of some 431 birds, mammals, frogs, reptiles and fish shows that genetic divergence between sister species ranges fairly evenly through the Pleistocene and Pliocene⁶⁹. Such analysis has been applied specifically to North American birds⁷⁰. On average, but with considerable variation and depending on species definition, speciation duration is about 1–2 Myr, and speciation appears to have been proceeding unhindered through the Quaternary period.

There is now clear evidence of extremely rapid differentiation in insects, lizards and fish in new habitats⁷¹. From molecular data the sympatric ecomorph species of sticklebacks, arctic charr and lake whitefish were probably formed in lakes as the ice retreated^{16,72}. On the other hand sister taxa of some species for which molecular data shows millions of years of divergence still hybridize, for example, toads and newts.

The current evidence indicates that although genomic divergence proceeds over millions of years, morphological, physiological and behavioural changes may produce subspecies and species at almost any time.

Many of these new local species will become extinct with the next major climatic shift, when their present location is covered by ice, tundra, savannah or desert, for example. By good fortune one may move and survive to found a more extensive new taxon. In analysing DNA divergences over the Plio-Pleistocene, we are observing lineages that have survived many ice ages. It is a species survival rate—a product of high speciation rate and periodic extinction. Mountain regions in the tropics have a greater range of life-supporting environments with greater continuity through climatic shifts, which would provide more opportunity for newly formed species to survive. This may partly explain greater species richness in the tropics.

Modern man and the ice ages

The ice ages can be expected to have also affected the evolution of man. Molecular genetic and archaeological studies are now being joined to elucidate the origin and spread of modern man worldwide. Most evidence supports an African origin followed by an expansion to other parts^{23,73–76}. The bones and tools of forms of *Homo erectus* are found in Africa and over Eurasia from 1.5 Myr ago. In Europe these produced Neanderthals from 300 kyr ago. DNA sequence from a fossil bone places their divergence about 600 kyr ago, with no evidence of similar sequences in present human data⁷⁷. Archaic *Homo sapiens* fossils and tools occur in Africa from 500 kyr ago, progressing to those of modern *H. sapiens* from 200 kyr to 100 kyr ago; a period coincident with the major ice ages.

A number of recent studies on molecular genetic diversity in present populations argue for an effective population size of about 10,000 in Africa during the Upper Pleistocene, through what is called a long bottleneck²³. These include mtDNA, Y chromosome, Alu insertion, human lymphocyte antigen (HLA), β -globin, the zinc-finger gene on the X chromosome (ZFX) intron and micro-satellite coalescent estimates. Mismatch distributions and neighbour-joining dendrograms suggest a population expansion ~100–50 kyr ago in Africa and 23 kyr ago in Europe⁷³. From ~130 kyr to 71 kyr (OIS5) ago, during and following the last interglacial, the climate was fairly warm with some reversals, and modern man entered the Levant⁷⁸. Around 71 kyr ago, the GRIP ice core record shows a very cold spell that is coincident with the volcanic eruption of Mount Toba in Sumatra, the largest anywhere for 450 Myr. It is argued by Ambrose⁷⁹ that this reduced the population of modern man to tropical refugia from which he expanded ~55 kyr ago. Significantly, Neanderthals who were probably better adapted for cold climates moved into the Levant from 71 kyr ago when modern man vacated it⁷⁸.

Modern man entered Europe from the Levant about 41 kyr ago with new technology and social structure, spreading across southern parts and replacing the Neanderthals gradually⁸⁰. The Neolithic agricultural revolution spread from the Near East around 9–5 kyr ago in the warm climate after the Younger Dryas cold reversal. This advance is correlated with clines in classical gene frequencies⁸¹. Artefacts show range changes between his arrival and this later spread, and man would have been pushed to southern Europe at the height of the ice age 24–20 kyr ago. Mitochondrial DNA data of modern populations^{82,83} show six principal lineages of this Upper Palaeolithic age, which have different divergences. The question being addressed now is how much of our genome originates from these earlier ice-age settlers: the contribution from the Neolithic revolution may be less than we thought. □

1. Bennett, K. J. *Evolution and Ecology: The Pace of Life* (Cambridge Univ. Press, Cambridge, 1997).

2. Williams, D., Dunkerley, D., DeDecker, P., Kershaw, P. & Chappell, M. *Quaternary Environments* (Arnold, London, 1998).

3. Webb, R. S., Rind, D. H., Lehman, S. J., Healy, R. J. & Sigman, D. The influence of ocean heat transport on the climate of the last Glacial Maximum. *Nature* **385**, 695–699 (1997).

4. Stauffer, B. Cornucopia of ice core results. *Nature* **399**, 412–413 (1999).
5. Reille, M., Andrieu, V., De Beaulieu, J. -L., Guenet, P. & Goery, C. A long pollen record from Lac du Bouchet, Massif Central, France: for the period ca. 325 to 100 ka BP (OIS 9c to OIS 5e). *Quat. Sci. Rev.* **17**, 1107–1123 (1998).
6. Behl, R. J. & Kennett, J. P. Brief interstadial events in the Santa Barbara basin, NE Pacific, during the past 60 kyr. *Nature* **379**, 243–246. (1996).
7. Schulz, A., von Rad, V. & Erlenkeuser, H. Correlation between Arabian Sea and Greenland climatic oscillations of the past 110,000 years. *Nature* **393**, 54–57 (1998).
8. Van Andel, T. H. & Tzedakis, P. C. Palaeolithic landscapes of Europe and environs, 150,000–25,000 years ago: an overview. *Quat. Sci. Rev.* **15**, 481–500 (1996).
9. Coope, G. R. The response of insect faunas to glacial-interglacial climatic fluctuations. *Phil. Trans. R. Soc. Lond. B* **344**, 19–26. (1994).
10. Rohling, R. J. *et al.* Magnitudes of sea-level lowstands of the past 500,000 years. *Nature* **394**, 162–165 (1998).
11. Colinvaux, P. A. An arid Amazon? *Trends in Ecology and Evolution* **12**, 318–319 (1997).
12. Hewitt, G. M. in *Speciation and its Consequences* (eds. Otte, D. & Endler, J.) 85–110 (Sinauer Associates, Sunderland, MA, 1989).
13. Hewitt, G. M. in *Hybrid zones and Evolutionary Process* (ed. Harrison, R. G.) 140–164 (Oxford Univ. Press, Oxford, 1993).
14. Ibrahim, K. M., Nichols, R. A. & Hewitt, G. M. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity* **77**, 282–291 (1996).
15. Hewitt, G. M. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.* **58**, 247–276 (1996).
16. Bernatchez, L. & Wilson, C. C. Comparative phylogeography of nearctic and paleartic fishes. *Mol. Ecol.* **7**, 431–452 (1998).
17. Gray, A. J. in *Past and Future Rapid Environmental Changes* NATO ASI Series, vol. 147 (eds Huntley B. *et al.*) 371–380 (Springer, Berlin, 1997).
18. Showers, W. B. in *Insect Life History Patterns* (eds Denno, R. & Dingle, H.) 97–111 (Springer, New York, 1981).
19. Wade, M. J. & Goodnight, C. J. The theories of Fisher and Wright in the context of metapopulations: when nature does many small experiments. *Evolution* **52**, 1537–1553 (1998).
20. Armbruster, P., Bradshaw, W. E. & Holzapfel, C. M. Effects of postglacial range expansion on allozyme and quantitative genetic variation of the pitcher-plant mosquito *Wyeomyia smithii*. *Evolution* **52**, 1697–1704 (1998).
21. Templeton, A. R. Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Mol. Ecol.* **7**, 381–397 (1998).
22. Hewitt, G. M. Post-glacial recolonization of European biota. *Biol. J. Linn. Soc.* **68**, 87–112 (1999).
23. Harpending, H. C. *et al.* Genetic traces of ancient demography. *Proc. Natl Acad. Sci. USA* **95**, 1961–1967 (1998).
24. Bandelt, H.-J., Forster, P. & Rohl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **16**, 37–48 (1999).
25. Taberlet, P., Fumagalli, L., Wust-Saucy, A. G. & Cossons, J.-F. Comparative phylogeography and postglacial colonization routes in Europe. *Mol. Ecol.* **7**, 453–464 (1998).
26. Cooper, S. J. B., Ibrahim, K. M. & Hewitt, G. M. Postglacial expansion and genome subdivision in the European grasshopper *Chorthippus parallelus*. *Mol. Ecol.* **4**, 49–60 (1995).
27. Flanagan, N. S., Mason, P. L., Gosalvez, J. & Hewitt, G. M. Chromosomal differentiation through an Alpine hybrid zone in the grasshopper *Chorthippus parallelus*. *Evol. Biol.* **12**, 577–585 (1999).
28. King, R. A. & Ferris, C. Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. *Mol. Ecol.* **4**, 95–103 (1998).
29. Santucci, F., Emerson, B. & Hewitt, G. M. Mitochondrial DNA phylogeography of European hedgehogs. *Mol. Ecol.* **7**, 1163–1172 (1998).
30. Ferris, C., King, R. A., Vainola, R. & Hewitt, G. M. Chloroplast DNA recognises three refugial sources of European oaks and shows independent eastern and western immigrations to Finland. *Heredity* **80**, 584–593 (1998).
31. Taberlet, P. & Bouvet, J. Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear (*Ursus arctos*) in Europe. *Proc. R. Soc. Lond. B* **255**, 195–200 (1994).
32. Jaarola, M., Tegelstrom, H. & Fredga, K. Colonization history in Fennoscandian rodents. *Biol. J. Linn. Soc.* **68**, 113–127 (1999).
33. Remington, C. L. Suture-zones of hybrid interaction between recently joined biotas. *Evol. Biol.* **2**, 321–428 (1968).
34. Avise, J. C. The history and purview of phylogeography: a personal reflection. *Mol. Ecol.* **7**, 371–379. (1998).
35. Avise, J. C. *et al.* Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. System.* **18**, 489–522 (1987).
36. Avise, J. C. in *Conservation Genetics* (eds Avise, J. C. & Hamrick, J. L.) 431–470 (Chapman & Hall, New York, 1996).
37. Walker, D. & Avise, J. C. Principles of phylogeography as illustrated by freshwater and terrestrial turtles in the southeastern United States. *Annu. Rev. Ecol. System.* **29**, 23–58 (1998).
38. Weisrock, D. W. & Janzen, F. J. Comparative molecular phylogeography of North American softshell turtles (*Apalone*): implications for regional and wide-scale historical evolutionary forces. *Mol. Phylogenetics Evol.* **14**, 152–164 (2000).
39. Soltis, D. E., Gitzendanner, M. A., Strenge, D. D. & Soltis, P. S. Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant System. Evol.* **206**, 353–373 (1997).
40. Conroy, C. J. & Cook, J. A. Phylogeography of a post-glacial colonizer: *Microtus longicaudus* (Rodentia: Muridae). *Mol. Ecol.* **9**, 165–175 (2000).
41. Zink, R. M. Comparative phylogeography in North American birds. *Evolution* **50**, 308–317 (1996).
42. Orange, D. I., Riddle, B. R. & Nickle, D. C. Phylogeography of a wide ranging desert lizard, *Gambelia wislizenii* (Crotaphytidae). *Copeia* **1999**, 267–273 (1999).
43. Ashworth, A. C. in *Past and Future Rapid Environmental Changes*. NATO ASI Series, vol. 147 (eds Huntley, B., *et al.*) 119–128 (Springer, Berlin, 1997).
44. Kidd, M. G. & Friesen, V. L. Analysis of mechanisms of microevolutionary change in *Cephus guillemoti* using patterns of control region variation. *Evolution* **52**, 1158–1168 (1998).
45. Wenink, P. W., Barker, A. J., Rosner, H. -U. & Tilanus, M. G. J. Global mitochondrial DNA phylogeography of holarctic breeding dunlins (*Calidris alpina*). *Evolution* **50**, 318–330 (1996).
46. Gravlund, P., Meldgaard, M., Paabo, S. & Arctander, P. Polyphyletic origin of the small-bodied, high-arctic subspecies of tundra reindeer (*Rangifer tarandus*). *Mol. Phylogenetics Evol.* **10**, 151–159 (1998).
47. Talbot, S. L. & Shields, S. L. Phylogeography of Brown Bears (*Ursos arctos*) of Alaska and paraphyly within the Ursidae. *Mol. Phylogenetics Evol.* **5**, 477–494 (1996).
48. Colbourne, J. K. *et al.* Phylogenetics and evolution of a circumpolar species complex (Cladocera: *Daphnia pulex*). *Biol. J. Linn. Soc.* **65**, 347–365 (1998).
49. Colinvaux, P. A., Oliviera, P. E., Moreno, J. E., Miller, M. C. & Bush, M. B. A long pollen record from lowland Amazonia: forest and cooling in glacial times. *Science* **274**, 85–88 (1996).
50. Schneider, C. J., Cunningham, M. & Moritz, C. Comparative phylogeography and the history of endemic vertebrates in the Wet Tropics rainforests of Australia. *Mol. Ecol.* **7**, 487–498 (1998).
51. Patton, J. L. & da Silva, M. N. F. in *Endless Forms: Species and Speciation*. (eds Howard, D. & Berlocher, S.) 202–213 (Oxford Univ. Press, Oxford, 1997).
52. Da Silva, M. N. F. & Patton, J. L. Molecular phylogeography and the evolution and conservation of Amazonian mammals. *Mol. Ecol.* **7**, 475–486 (1998).
53. Bermingham, E. & Martin, A. P. Comparative mtDNA phylogeography of neotropical freshwater fishes: testing shared history to infer the evolutionary landscape of lower Central America. *Mol. Ecol.* **7**, 499–517 (1998).
54. Arctander, P., Johansen, C. & Coutellec-Vreto, M. A. Phylogeography of three closely related African bovids (tribe Alcelaphini). *Mol. Biol. Evol.* **16**, 1724–1739 (2000).
55. Fjeldsa, J. Geographical patterns for relict and young species of birds in Africa and South America and implications for conservation priorities. *Biodivers. Conserv.* **3**, 207–226 (1994).
56. Fjeldsa, J. & Lovett, J. C. Geographical patterns of old and young species in African forest biota: the significance of specific montane areas as evolutionary centres. *Biodivers. Conserv.* **6**, 323–344 (1997).
57. Roy, M. S. Recent diversification in African greenbul (Pycnonotidae: *Andropadus*) supports a montane speciation model. *Proc. R. Soc. Lond. B* **264**, 1337–1344 (1997).
58. Garcia-Moreno, J., Arctander, P. & Fjeldsa, J. A case of rapid diversification in the Neotropics: Phylogenetic relationships among *Cranioleuca* spinetails (Aves, Furnariidae). *Mol. Phylogenetics Evol.* **12**, 273–281 (1999).
59. Ruedi, M. & Fumagalli, L. Genetic structure of Gymnures (genus *Hylomys*: Erinaceidae) on continental islands of Southeast Asia: historical effects of fragmentation. *J. Zool. System. Evol. Res.* **34**, 153–162 (1996).
60. Dodson, J. J., Colombani, F. & Ng, P. K. L. Phylogeographic structure in mitochondrial DNA of a Southeast Asian freshwater fish, *Hemibagrus nemurus* (Siluroidei: Bagridae) and Pleistocene sea-level changes on the Sunda shelf. *Mol. Ecol.* **4**, 331–346 (1995).
61. Surridge, A. K., Timmins, R. J., Hewitt, G. M. & Bell, D. J. Striped rabbits in Southeast Asia. *Nature* **400**, 726 (1999).
62. Hisheh, S., Westerman, M. & Schmitt, L. H. Biogeography of the Indonesian archipelago: mitochondrial DNA variation in the fruitbat, *Eonycteris spelaea*. *Biol. J. Linn. Soc.* **65**, 329–345 (1998).
63. Palumbi, S. R. Molecular biogeography of the Pacific. *Coral Reefs* **16**, 547–552 (1997).
64. Williams, S. T. & Benzie, J. A. H. Evidence of a biogeographic break between populations of a high dispersal starfish: congruent regions within the Indo-West Pacific defined by color morphs, mt DNA, and allozyme data. *Evolution* **52**, 87–99 (1998).
65. Chenoweth, S. F., Hughes, J. M., Keenan, C. P. & Lavery, S. When oceans meet: a teleost shows secondary intergradation at an Indian-Pacific interface. *Proc. R. Soc. Lond. B* **265**, 415–420 (1998).
66. Duke, N. C., Benzie, J. A. H., Goodall, J. A. & Ballment, E. R. Genetic structure and evolution of species of the mangrove genus *Avicennia* (Avicenniaceae) in the Indo-West Pacific. *Evolution* **52**, 1612–1626 (1998).
67. Bowen, B. W. & Grant, W. S. Phylogeography of the sardines (*Sardinops* spp.): assessing biogeographic models and population histories in temperate upwelling zones. *Evolution* **51**, 1601–1610 (1997).
68. Hauser, L. & Ward, R. in *Advances in Molecular Ecology* (ed. Cavalho, G.) 191–223 (IOS Press, Amsterdam, 1998).
69. Avise, J. C., Walker, D. & Johns, G. C. Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proc. R. Soc. Lond. B* **265**, 1707–1712 (1998).
70. Klicka, J. & Zink, R. M. Pleistocene effects on North American songbird evolution. *Proc. R. Soc. Lond. B* **266**, 695–700 (1999).
71. Orr, M. R. & Smith, T. B. Ecology and Speciation. *Trends Ecol. Evol.* **13**, 502–506 (1998).
72. Schluter, D. in *Endless Forms: Species and Speciation*. (eds Howard, D. & Berlocher, S.) 114–129. (Oxford Univ. Press, Oxford, 1998).
73. Jorde, L. B., Bamshad, M. & Rogers, A. R. Using mitochondrial and nuclear DNA markers to reconstruct human evolution. *BioEssays* **20**, 126–136 (1998).
74. Templeton, A. R. Out of Africa? What do genes tell us? *Curr. Opin. Genet. Dev.* **7**, 841–847 (1997).
75. Stringer, C. B. *African Exodus: The Origins of Modern Humanity* (Jonathan Cape, London, 1996).
76. Hublin, J.-J. in *Neandertals and Modern Humans in Western Asia* (eds Akazawa, T. *et al.*) (Plenum, New York, 1998).
77. Krings, M., Stone, A., Schmitz, R. W. & Krainitzki, H. Neandertal DNA sequences and the origins of modern humans. *Cell* **90**, 19–30 (1997).
78. Tchernov, E. in *Neandertals and Modern Humans in Western Asia* (eds Akazawa, T. *et al.*) (Plenum, New York, 1998).
79. Ambrose, S. H. Late Pleistocene human population bottlenecks, volcanic winter, and differentiation of modern humans. *J. Hum. Evol.* **34**, 623–651 (1998).
80. Mellars, P. in *Neandertals and Modern Humans in Western Asia* (eds Akazawa, T. *et al.*) (Plenum, New York, 1998).
81. Cavalli-Sforza, L. L., Menozzi, P. & Piazza, A. *The History and Geography of Human Genes* (Princeton Univ. Press, Princeton NJ, 1994).
82. Richards, M. *et al.* Paleolithic and neolithic lineages in the European mitochondrial gene pool. *Am. Hum. Genet.* **59**, 185–203 (1996).
83. Richards, M., MacAulay, V., Bandelt, H. -J. & Sykes, B. Phylogeography of mitochondrial DNA in western Europe. *Ann. Hum. Genet.* **62**, 241–260 (1998).
84. Li, W. H. *Molecular Evolution*, (Sinauer Associates, Sunderland, MA, 1997).
85. Parsons, T. J. & Holland, M. M. Mitochondrial mutation rate revisited: hotspots and polymorphism. *Nature Genet.* **18**, 110 (1998).
86. Schug, M. D., Mackay, T. F. C. & Aquadro, C. F. Low mutation rates of microsatellite loci in *Drosophila melanogaster*. *Nature Genet.* **15**, 99–102 (1997).

Correspondence and requests for materials should be addressed to G.H. (e-mail: g.hewitt@uea.ac.uk).

Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export

Vassilis Koronakis^{*†}, Andrew Sharff^{†‡}, Eva Koronakis^{*}, Ben Luisi[‡] & Colin Hughes^{*}

^{*} Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK

[‡] Crystallography & Biocomputing Unit, Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, UK

[†] These authors contributed equally to this work

Diverse molecules, from small antibacterial drugs to large protein toxins, are exported directly across both cell membranes of Gram-negative bacteria. This export is brought about by the reversible interaction of substrate-specific inner-membrane proteins with an outer-membrane protein of the TolC family, thus bypassing the intervening periplasm. Here we report the 2.1-Å crystal structure of TolC from *Escherichia coli*, revealing a distinctive and previously unknown fold. Three TolC protomers assemble to form a continuous, solvent-accessible conduit—a ‘channel-tunnel’ over 140 Å long that spans both the outer membrane and periplasmic space. The periplasmic or proximal end of the tunnel is sealed by sets of coiled helices. We suggest these could be untwisted by an allosteric mechanism, mediated by protein–protein interactions, to open the tunnel. The structure provides an explanation of how the cell cytosol is connected to the external environment during export, and suggests a general mechanism for the action of bacterial efflux pumps.

Export of a wide range of molecules across the cell envelope of Gram-negative bacteria is achieved in a single, energy-coupled step, resulting in direct passage across both inner and outer cell membranes and the intervening periplasmic space. The underlying mechanism is determined by the direct interaction of an inner membrane translocase of two proteins, which provides substrate

specificity and energy, with a third in the outer membrane—a protein of the TolC family^{1,2}. When engaged by its export substrate, the translocase recruits TolC to form the active export complex containing all three proteins and the substrate¹. This assembly is transient: when the substrate has been exported, the machinery disengages and the components revert to their separate inner and

Table 1 Summary of data collection and refinement statistics

Compound	Native	Mercury	Oxidized SeMet	
Space group	<i>R</i> 3	<i>R</i> 3	<i>R</i> 3	<i>R</i> 3
Unit cell <i>a</i> (Å)	265.50	262.56	265.11	265.05
<i>c</i> (Å)	95.50	95.53	96.00	95.96
Beamline	SRS 9.6	ESRF BM14	ESRF BM14	APS ID19
Temperature (K)	100	100	100	100
No. of crystals used	1	1	1	1
Wavelength (Å)	0.8700	0.9793	0.9788	0.9790
Max. resolution (Å)	2.50	4.00	2.95	2.95
Unique reflections	86838	18509	52877	52902
Completeness (%)	99.9	98.9	99.9	99.9
Redundancy	9.0	4.0	5.7	5.7
<i>I</i> / <i>σ</i> [*]	5.1 (2.3)	8.0 (4.7)	4.5 (4.2)	4.1 (4.2)
<i>R</i> _{sym} (%) [†]	8.8 (18.5)	6.2 (12.2)	7.3 (13.4)	7.1 (13.9)
<i>R</i> _{ano} (%) [†]		6.3 (8.7)	5.7 (7.4)	5.7 (8.7)
Anomalous phasing power		1.46‡	1.1	2.1
<i>R</i> _{culis} iso/ano§		0.95/0.95°	−0.91	0.57/0.73
Figure of merit				0.57/0.86
After SHARP refinement		0.156‡		0.450
After solvent flattening		0.586‡		0.840
Refinement				
Resolution range				20.0–2.1
<i>R</i> factor (%)				20.8
<i>R</i> _{free} (%)°				25.7
Ordered water molecules				1,508
R.m.s. deviation from ideal geometry				
Bond lengths (Å)				0.018
Bond angles (°)				1.3
Average <i>B</i> factor (Å ²)				49.88
R.m.s. deviation of monomers (Å)				0.18 ± 0.02

^{*} Figures in parentheses indicate values in the last resolution shell.

[†] $R_{\text{sym}} = \sum |I_{\text{avg}} - I| / \sum I$, and $R_{\text{ano}} = \sum |I_{+} - I_{-}| / \sum |I_{+} + I_{-}|$.

[‡] Data only used to 5.0 Å resolution.

[§] $R_{\text{culis}} = \sum ||F_{\text{PH}} \pm F_{\text{P}}| - |F_{\text{H}}(\text{calc})|| / \sum |F_{\text{PH}}|$.

^{||} $R = \sum |F_{\text{obs}} - F_{\text{calc}}| / \sum F_{\text{obs}}$ × *R*_{free} is calculated from about 5% of randomly chosen reflections, which are not used in refinement.

outer membrane resting states¹. This 'shuttle' mechanism does not generate periplasmic export intermediates and is fundamentally distinct from processes that determine step-wise translocation across two membranes^{3–5}. TolC and its homologues interact with different translocase complexes, depending on the transport substrate, so that combinatorial assembly provides functional diversity. This system exports large proteins (type I secretion), including enzymes and toxins such as cyclolysin (relative molecular mass, M_r 170,000)^{1,2,6–8}, and small noxious agents, such as detergents, solvents, heavy metals and antibiotics (multidrug efflux)^{2,9,10}. TolC is therefore an important component in the determination of both pathogenicity and multidrug resistance, and loss of TolC or one of its homologues reduces bacterial survival and attenuates virulence¹¹. This suggests that these proteins may be potential chemotherapeutic targets.

Biochemical and electron microscopy studies have shown that TolC is a trimeric outer membrane protein¹², but a simple porin-like model would not readily suggest a means for interaction with the inner membrane translocase during transport. The detailed three-dimensional structure of TolC that we present here provides insights into the protein's function and the mechanism of export. It reveals a new architectural design, the α -helical barrel, which forms a tunnel through the periplasm, anchored by a contiguous outer membrane β -barrel. A stereochemical model is proposed for the opening of this structure.

Structure determination

The mature 471-residue TolC protein was detergent-extracted from

E. coli membranes and purified to homogeneity. To obtain crystals, it was necessary to treat TolC with V8 protease, which removed the carboxy-terminal 43 residues. Subsequent assays showed that truncation does not affect function (E.K., unpublished data). Crystals of the protease-truncated TolC grew in space group $R3$ with cell dimensions $a = b = 265 \text{ \AA}$, $c = 95 \text{ \AA}$, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$ (in the hexagonal setting). In preparation for data collection, crystals were frozen at 100 K after cryoprotection. The structure was solved by multiple wavelength anomalous dispersion (MAD) using oxidized selenomethionine (SeMet) derivatives¹³.

Data extending to 3.0 $\text{\AA}$ and 2.1 $\text{\AA}$ were collected from separate crystals around the Se K absorption edge, and these were used to solve and refine the structure. Data were also collected from crystals of native (non-derivatized) protein, without oxidation, at 2.5 $\text{\AA}$, and a model was refined against these data. Comparison of the native and oxidized SeMet models indicates no significant structural differences. The crystals contain one trimer of M_r 141K in the asymmetric unit and have a solvent content of 70%, which is high for protein crystals. The final model has been refined at 2.1 $\text{\AA}$ with an R factor of 21% and R_{free} of 25%. The diffraction data quality and refinement statistics are summarized in Table 1 (see Supplementary Information).

Overall architecture

TolC is assembled as a trimer of 428-residue protomers (Fig. 1a). The appearance of the trimer is that of a cannon, with a long axis measuring 140 $\text{\AA}$. For nearly 100 $\text{\AA}$ the body forms a uniform cylinder of about 35 $\text{\AA}$ internal diameter, in rough agreement with

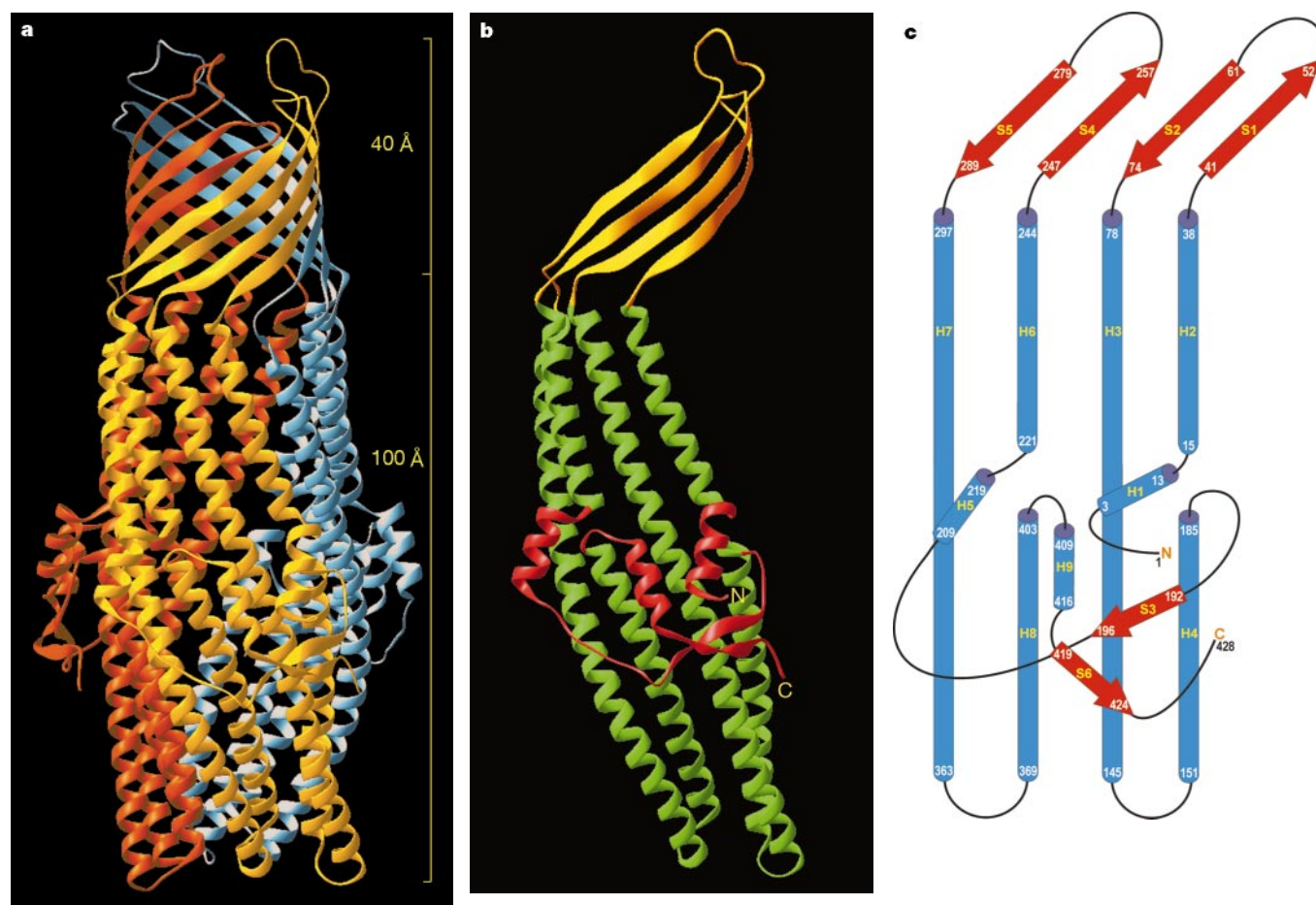


Figure 1 The overall architecture of TolC. **a**, C α trace of TolC. The protomers are individually coloured. The molecular threefold axis is aligned vertically, normal to the plane of the outer membrane. The β -barrel is at the top (distal) end, and the α -helical domain is at the bottom (proximal) end. Figure generated with RIBBONS³⁹. **b**, C α trace of a single

protomer. The β -barrel domain is yellow, the α -helical barrel domain is green and the equatorial domain is red. **c**, Topology diagram of the protomer. Secondary-structure elements are indicated: helices (H) in blue, strands (S) in red. The structural repeat comprises the sets (H1, H2, S1, S2, H3, H4) and (H5, H6, S4, S5, H7, H8).

previous electron microscopy studies¹². The distal (upper) end of the structure is open (see Supplementary Information) and provides a wide solvent access while the proximal (lower) end is tapered to a virtual close. This explains the previous observation that although TolC forms an ion-permeable channel when reconstituted into lipid bilayers, its conductance is unusually low¹⁴. The body has a large interior cavity that is mostly solvent-filled with a volume of roughly 43,000 Å³, making it one of the largest known in a protein structure.

The TolC molecule can be partitioned into a β -domain, an α -helical domain, and a mixed α/β -domain (Fig. 1b). The peptide chain of each protomer weaves up and down the long axis four times, and passes from β -strands (S1, S2, S4, S5), at the distal end of the structure, into α -helices, two of which (H3 and H7) extend to the proximal end (Fig. 1c). The α/β -domain, which we will refer to as the equatorial domain, is made up of strands S3 and S6 and helices H1, H5 and H9 and forms a 'strap' around the mid-section of the helical barrel.

In the trimer, the strands of the β -domain associate in an antiparallel orientation to form a 12-stranded β -barrel that is right-twisted (that is, the strands have a positive inclination with respect to the molecular threefold axis). The α -helices forming the main body of the structure also form a 12-stranded antiparallel

barrel, but in contrast to the β -barrel, this α -helical barrel is left-twisted. The opposite twist of the two supersecondary elements originates from the propensity of β -sheets to twist and curve clockwise in globular proteins and of coiled-coil helices to supercoil anticlockwise (for certain sequence repeat patterns¹⁵). The transitions from right-twisted β -barrel into left-twisted α -helical barrel are accommodated through abrupt turns in proline-containing interdomain linkers. The importance of these prolines is underscored by their conservation in the family of bacterial outer membrane efflux proteins. The α -helical barrel is constructed from two types of helix. Long helices (H3 and H7), comprising 67 residues, traverse the entire length from the lower end of the β -barrel to the proximal end of the molecule. The second type are shorter helices of 23 (H2 and H6) and 34 (H4 and H8) residues, respectively. The pairs of the shorter helices stack end to end to form pseudocontinuous helices (H2 and H4; H6 and H8) that, like the long helices, traverse the length of the helical barrel. The equatorial domain is connected to the helical barrel at the junction of these short helices.

The TolC structure comprises an approximate structural repeat (H1–H2–S1–S2–H3–H4 and H5–H6–S4–S5–H7–H8), which is evident in the schematic representation of the protomer (Fig. 1c). The structural repeat corresponds to a repeat in the primary

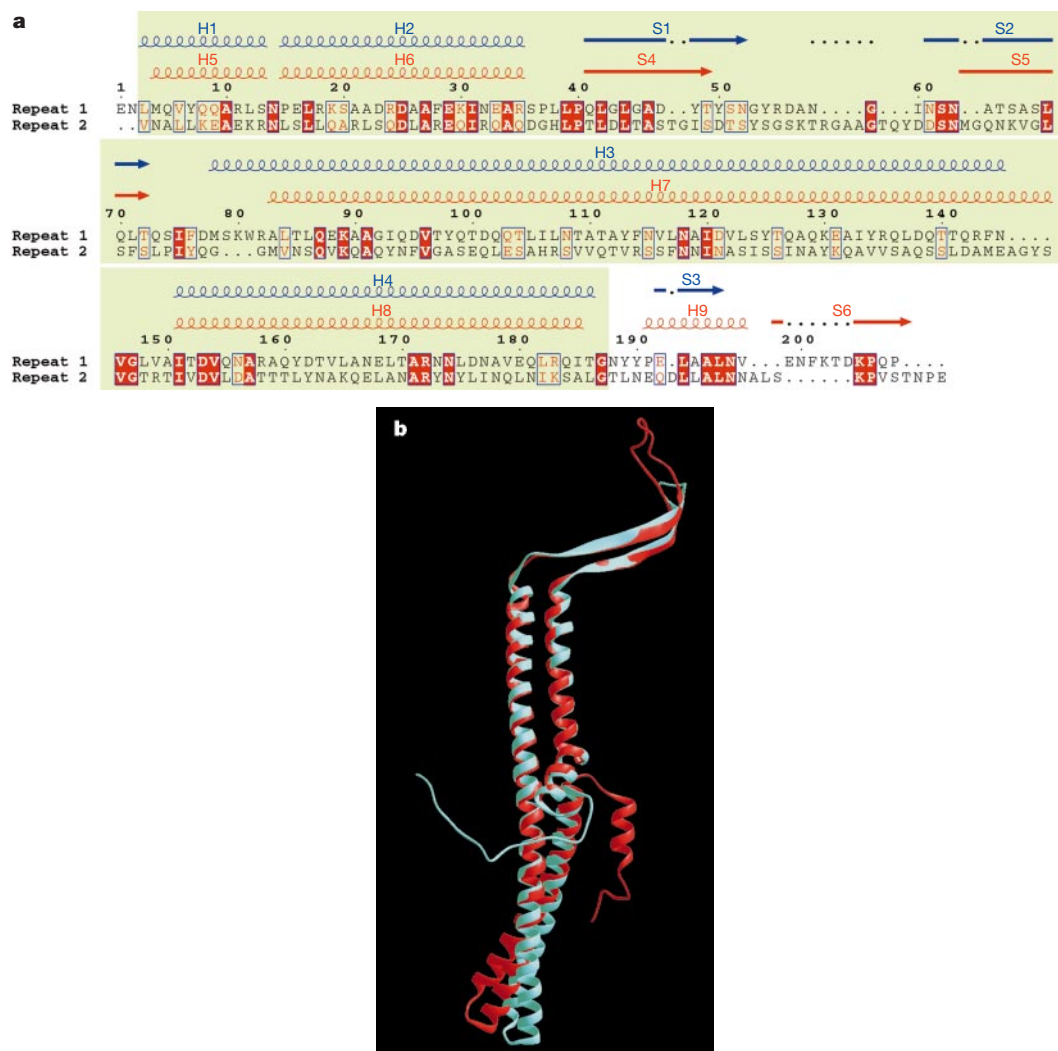


Figure 2 Structural repeat in TolC. **a**, Amino-acid sequence of TolC, annotated with the elements of secondary structure. The sequence has been internally aligned to highlight the structural repeat (shaded in light green). The secondary structure for the N-terminal

repeat is coloured in blue, the C-terminal repeat in red. **b**, Superimposed C α trace of the structural repeats. Repeat 1 is shown in blue, repeat 2 in red.

sequence (Fig. 2a, b), previously noted in the family of bacterial outer membrane efflux proteins¹⁶. These observations suggest that TolC and its relatives evolved by gene duplication from a common ancestor, which may have functioned as a hexamer.

The β -barrel domain

To date, all of the bacterial outer membrane proteins (OMPs) that have been characterized structurally are β -barrels of 8 to 22 strands, which are inserted in the membrane^{17–24}. As noted in these OMPs, aromatic residues, particularly tyrosine and phenylalanine, cluster in a ring around the base of the TolC β -barrel, delimiting the position of the inner edge of the lipid bilayer²⁵. This clearly indicates that the β -barrel of TolC is similarly located in the outer membrane. The β -barrel of TolC is assembled from three protomers, with each protomer contributing four β -strands. This is so far unique among bacterial OMPs, whose barrels are typically formed from a single monomer. Each of the strands have between 10 and 13 residues, which is common in OMPs. The β -strands must both curve and twist to attain their superhelical path, a movement analogous to the right-handed twisting motion usually associated with the strands in β -sheets in globular proteins. To accommodate the curvature, small or unbranched side chains tend to lie periodically on the inside of the barrel, where they allow the required close packing. Out of phase with the interior repeats, bulkier, β -branched residues lie on the outside of the barrel, where they accommodate the gaps caused by the curving peptide trajectory.

The interiors of the OMPs are often partially or completely occluded^{17,22–25}. In contrast, the β -barrel of TolC is wide open and fully accessible to solvent. The loops at the top end of the β -strands of TolC may close over the mouth as a partial external 'lid', but they have comparatively high atomic displacement parameters, indicating that they probably have conformational mobility.

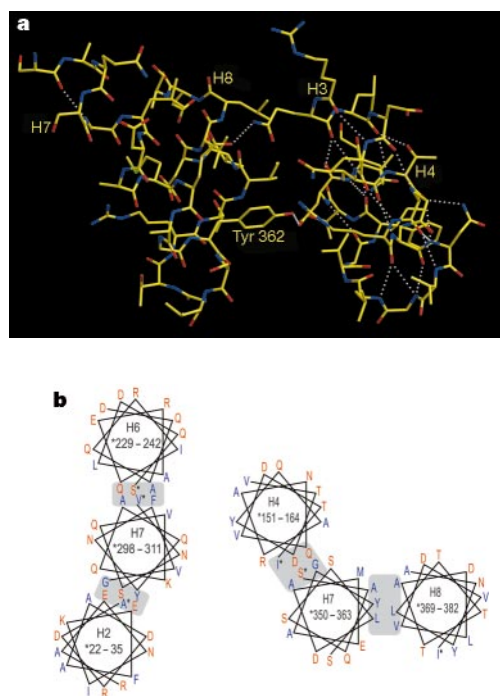


Figure 3 Interfacial contacts in the α -helical barrel. **a**, View of selected contacts in the interface between H3/H4 and H7/H8. The contact made by Tyr 362 is well conserved in the TolC family. We propose that this contact is broken in the open state of the channel. **b**, Helical wheel representations summarizing the interhelical contacts in different parts of the α -helical barrel. The register of lateral contacts between neighbouring helices at the top end of the helical barrel (left) can be maintained in a superhelix, whereas those at the proximal end, between sets of coiled coils (right), can only occur locally. Asterisks indicate the N-terminal residue of each helical segment.

The α -helical domain and coiled-coil interactions

The α -helical barrel is contiguous with the β -barrel and extends into the periplasm. This structure has not been observed before, and it has distinctive features. The upper section of the helical domain (near the β -barrel) is virtually uniform in diameter, and the helices are inclined by -20° relative to the molecular threefold axis. It may not be readily apparent that these helices are coiled coils, but, because they follow an inclined trajectory on the surface of a regular cylinder, they must both curve and twist in space. The TolC helices have a left-handed superhelical twist throughout the entire α -helical barrel, but tend to untwist at the distal end compared with conventional coiled coils. There are no abrupt kinks or bends in either of the long helices, and the continuous change in superhelical trajectory is distributed in small deviations in the local helical twist, as found in other coiled coils²⁶.

The trajectories of the long helices (H3 and H7) remain uniform preceding the equatorial domain, but depart from the cylindrical surface thereafter to form the taper at the proximal end (Fig. 1b). Here, an inner pair of helices (H7 and H8) form a conventional antiparallel coiled-coil. In contrast, an outer pair (H3 and H4) comprise a straight helix (H3) around which its partner (H4) is coiling. The coiled coils also contact each other (Fig. 3a). Given the sequence similarity of the inner and outer pairs of helices, it is striking that they pack so differently when they could pack in a quasi-equivalent manner. This point may have some bearing on the mechanism of channel opening.

Coiled coils are stabilized by an intermeshing of side chains that is commonly referred to as 'knobs-into-holes' packing²⁷. A small aliphatic side chain of one helix fits into a small concavity formed by a ring of four small side chains on the second helix. Small non-polar side chains repeat with a characteristic pattern (a-b-c-d-e-f-g) where residues at positions a and d lie in the hydrophobic core of the helical interface. In TolC, the helices at the distal end of the barrel pack laterally with two neighbours and so form two separate interfaces (Fig. 3b). This means that each helix must have two

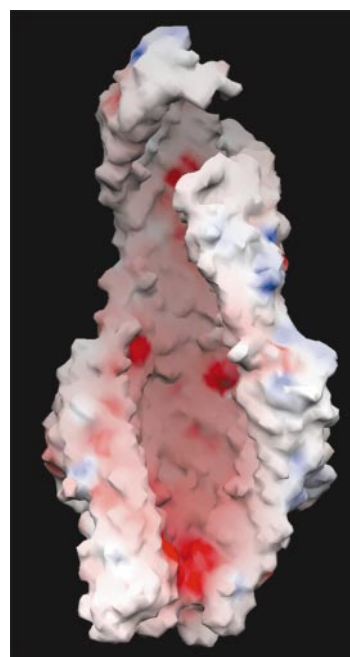


Figure 4 Surface representation of the TolC interior, illustrating the charge distribution. One protomer has been removed to expose the interior view. Charges are shown in red for electronegative and blue for electropositive; white is nonpolar. Figure generated with GRASP⁴⁰.

separate repeats that are out of phase by roughly half a helical turn so that they lie on different exposed faces. There is therefore an approximate equivalence between positions a and c and between d and f. In conventional coiled coils, the register of interfaces is achieved through superhelical twist, but in TolC the superhelical twist is smaller, with the result that the helices can lie on a cylindrical surface. A noteworthy feature of the cylindrical portion of the barrel is that bulkier side chains tend to lie on the exterior face of the barrel, and their steric bulk might accommodate the curvature of the barrel.

The equatorial domain

Unlike other parts of the TolC structure, the equatorial domain has no pseudosymmetry. The helices and strands of this mixed α/β structure pack against the helices of the helical barrel, and its fold is consolidated through numerous hydrogen bonds and van der Waals contacts between side chains. The C-terminal 43 residues that were proteolytically removed to prepare the crystals would extend from the equatorial domain.

Properties of the exposed surfaces

The exterior of the β -barrel is in direct contact with the aliphatic chains of the lipid bilayer in the outer membrane and is largely non-polar, as expected. The α -helical barrel and equatorial domains present a mixture of predominately nonpolar and isolated electro-

negative and electropositive patches. The surface of the interior cavity from the top of the structure down to the equatorial domain is also predominately nonpolar. On moving toward the proximal end of the α -helical barrel, however, the interior surface becomes increasingly electronegative (Fig. 4). This may have implications for the transport mechanism, as all substrates entering the TolC tunnel will encounter this striking electronegative surface. TolC and its homologues transport molecules with a wide variation in charge. It follows that a pulse of cations early in transport might favour entry of acidic and hydrophobic substrates into the tunnel, whereas a late pulse might catalyse the release of basic molecules.

A proposal for the transport mechanism

TolC is central to the export of diverse compounds, from small molecules to large proteins, in each case interacting with a specific inner membrane translocase. It is apparent that during transport, the tapered proximal end of the trimer must open and, in the case of proteins, the transport substrate must be at least partially unfolded. The simplest opening mechanism envisages that the inner pair of coiled coils rotates around its neighbouring partner to dilate the entrance. A clue as to the nature of the opening is provided by the structural repeat. The inner and outer sets of coiled coils (H7/H8 and H3/H4, respectively) have similar sequences, which means that they could in principle make similar interhelical interactions. The two sets of coiled coils differ only by small changes in superhelical twist, as shown in Fig. 2b, where we have superimposed the two structural repeats. Hence, the inner pair could re-pack to become congruent with the outer pair by an economical, untwisting movement. Comparable untwisting motions of coiled coils have been noted in other allosteric proteins²⁸. The proposal shares some similarity with the pore-opening mechanism proposed for the transmembrane helices of the acetylcholine receptor²⁹. Using the same reference frame as the superposition of Fig. 2b, the TolC open state was modelled by replacing the inner set of coiled coils with the outer set (Fig. 5). The uncoiling movement of the inner set of coiled coils could open the tunnel by as much as 30 Å.

We propose that the uncoiling movement is triggered by protein–protein interaction, that is, on recruitment of TolC by the inner membrane translocase. The equatorial domain is one possible recognition site for such interaction. Its strands and helices pack against the inner set of coiled coils, and any change in this relationship, induced by interactions with a partner protein, might activate an allosteric transition in the coiled coils. In the modelled open state, the inner coiled coil loses its contacts with the outer coil (for example, H3/H4 packing against H7/H8, see Fig. 3a), but these could be replaced by contacts with the bound translocase. It is noteworthy that the translocase components interacting with TolC¹ have sequence signatures for a coiled-coil motif. Perhaps in the open complex this coiled-coil domain re-packs against the unwound coiled coils at the proximal end of TolC. As the TolC protomer has approximate twofold symmetry, the open structure would have quasi-sixfold symmetry (Fig. 5), and could match the trimeric or hexameric partner in the inner membrane (it has been shown that the translocase partner interacting with TolC is at least trimeric¹).

The structure of the TolC channel tunnel reveals a mechanism to facilitate the direct passage of substrates across two membranes and the intervening periplasmic space. During transport, TolC is recruited by an energized inner membrane translocase in response to substrate engagement, providing substrate specificity and regulated access to the tunnel. This series of events establishes a continuous conduit that transiently connects the cell cytosol to the external environment through a completely proteinaceous machinery. This may have relevance beyond prokaryotic cell function, for example in the direct import of proteins across the two membranes of mitochondria³⁰. The sequence similarity of TolC family members suggests that the principal structural features are

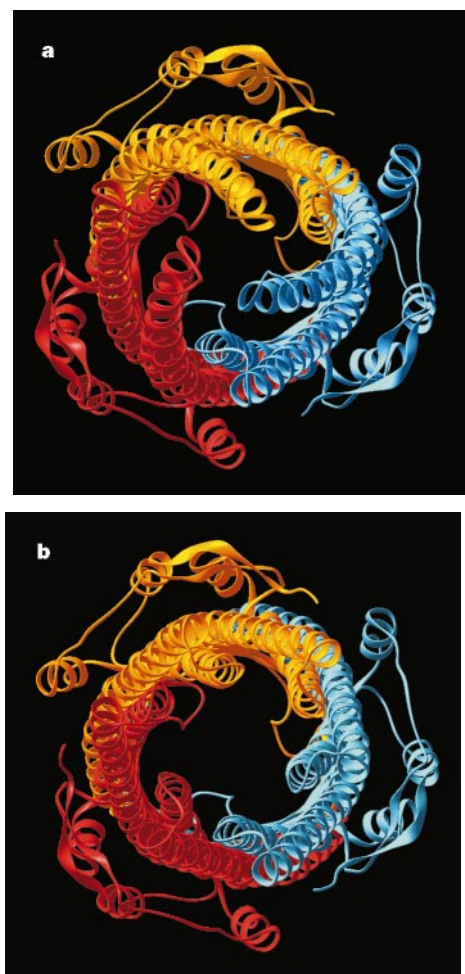


Figure 5 Exterior view of the proximal end of the α -helical barrel, approximately down the threefold symmetry axis. **a**, Crystal structure highlighting the coiled coils closing the proximal end of the tunnel. **b**, 'Open-state' model illustrating how the channel may be opened.

conserved, and that the allosteric mechanism proposed may be general to bacterial efflux pumps effecting transport of proteins, antibiotics and other molecules. □

Methods

Protein expression, purification and crystallization

The TolC protein was overexpressed in *E. coli* strain BL21(DE3), purified following detergent extraction as described¹² and treated with V8 protease (500 units per 30 mg protein, 24 h at 30 °C). This cleaved only once, removing the C-terminal 43 residues (a recombinant truncated TolC lacking these 43 residues retained wild-type TolC protein export and bile-salt resistance functions when expressed in *E. coli*). The cleaved protein was re-purified by anion-exchange (Q sepharose). Protein crystals were grown by vapour diffusion from hanging drops at 25 °C. The concentrations of reagents in the drops were a 0.6% detergent mixture of (*n*-dodecyl- β -D-glucopyranoside, *n*-hexyl- β -D-glucopyranoside, *n*-heptyl- β -D-glucopyranoside and *n*-octyl- β -D-glucopyranoside), 1.5% 1,2,3-heptanetriol, 7% polyethylene glycol 2000 monomethyl ether, 10% polyethylene glycol 400, 10 mM NaCl, 20 mM MgCl₂, 20 mM Tris, pH 7.4, and a protein concentration of 10–15 mg ml⁻¹. The drops were equilibrated against 12.5% polyethylene glycol 2000 monomethyl ether, 400 mM NaCl, 20 mM Tris, pH 7.4. Crystals usually appeared within 24–48 h. Initial screens with native TolC for heavy-atom derivatives were unsuccessful; however it was found that the addition of 0.5 mM potassium gold cyanide to the droplets slowed down nucleation and crystal growth, resulting in larger and much more strongly diffracting crystals. There was no evidence of bound gold in the crystals, and gold cyanide was routinely added to all subsequent crystallizations. Crystals of native TolC and derivatives were stabilized in 7% polyethylene glycol 2000 monomethyl ether, 10% polyethylene glycol 200, 15% polyethylene glycol 400, 1.5% 1,2,3-heptanetriol, 0.4% detergent mixture, 0.25 M NaCl, 20 mM MgCl₂, 20 mM Tris pH 7.4 and rapidly frozen at 100K for storage and data collection.

Data collection and structure determination

A series of 10 Ser→Cys point mutants were prepared and pretreated with a variety of mercury compounds before crystallization. Of those that gave usable crystals, only one (S402C) treated with mercury ethyl acetate showed some degree of derivatization. The crystals diffracted weakly and, although 4.0 Å data were collected on station BM14 at the ESRF (Grenoble), near the mercury LIII-edge, the data were useful only to 5.0 Å. Peaks were readily apparent in an anomalous Patterson map, corresponding to three equidistant mercury atoms per asymmetric unit, but the occupancy was very low (~0.15 per site) and the derivative contributed little to the overall phasing power. It was, however, useful at early stages of model building for corroborating the trace and determining the non-crystallographic symmetry operators.

A SeMet derivative was prepared after re-transforming into the met⁻ *E. coli* B834(DE3). Initial MAD data collections were unsuccessful, and we noted that the fluorescence spectra of the SeMet crystals were broad and lacked a characteristic 'white line' at the absorption peak, suggesting heterogeneity in the oxidation state of the selenium. Oxidation of the SeMet residues in TolC by pretreatment of the protein with 0.1% hydrogen peroxide before crystallization resulted in a sharpened fluorescence spectrum and significantly enhanced the amplitude of the fluorescence signal¹³. MAD data were collected from the oxidized SeMet TolC crystals at station BM14 at ESRF and station ID19 at the APS (Chicago) to 3.0 Å and 2.1 Å, respectively. The data were processed with HKL2000 (ref. 31).

Fourteen of the expected fifteen selenium sites were found from the 3.0 Å MAD data using SOLVE³². The final site was located from difference Fourier maps. The selenium parameters were refined and phases calculated using SHARP³³. Phases were improved by density modification, including threefold averaging through matrices defined by the heavy-metal sites, using DM^{34,35}. The initial figure of merit was 0.84 at 3.0 Å. The map was of sufficient quality to trace the entire backbone and assign side chains using the Se and Hg sites as landmarks. Subsequently, the model resolution was extended to 2.1 Å against data collected at the APS. The model was refined with CNS³⁶, applying strict non-crystallographic symmetry constraints, and improved using wARP³⁷ to model solvent. The final stages of refinement were carried out using BUSTER³⁸, which modifies the phases by a maximum likelihood algorithm. The resulting maps were substantially improved and helped to improve the clarity of the density in the poorly ordered loop regions of the β -barrel domain. The non-crystallographic symmetry restraints were relaxed during these later stages and the final cycle was carried out with no restraints. We observe no clear indication of localized lipid or detergent around TolC. The model was also refined against the 2.5 Å native data (collected at station 9.6, Daresbury). The two refined structures show that treatment with peroxide did not cause any detectable structural changes in the protein.

Received 14 March; accepted 4 May 2000.

1. Thanabalu, T., Koronakis, E., Hughes, C. & Koronakis, V. Substrate-induced assembly of a contiguous channel for protein export from *E. coli*: reversible bridging of an inner-membrane translocase to an outer membrane exit pore. *EMBO J.* **17**, 6487–6496 (1998).
2. Paulsen, I. T., Park, J. H., Choi, P. S. & Saier, M. H. A family of Gram-negative bacterial outer membrane factors that function in the export of proteins, carbohydrates, drugs and heavy metals. *FEBS Microbiol. Lett.* **156**, 1–8 (1997).
3. Koronakis, V., Koronakis, E. & Hughes, C. Isolation and characterisation of the C-terminal signal directing export of *E. coli* hemolysin across both bacterial membranes. *EMBO J.* **8**, 595–605 (1989).
4. Koronakis, V., Hughes, C. & Koronakis, E. Energetically distinct early and late stages of HlyB/HlyD-dependent secretion across both *Escherichia coli* membranes. *EMBO J.* **10**, 3263–3272 (1991).
5. Duong, F., Eichler, J., Price, A., Leonard, M. R. & Wickner, W. Biogenesis of the Gram-negative bacterial envelope. *Cell* **91**, 567–573 (1997).

6. Glaser, P., Sakamoto, H., Bellalou, J., Ullmann, A. & Danchin, A. Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-hemolysin bifunctional protein of *Bordetella pertussis*. *EMBO J.* **7**, 3997–4004 (1988).
7. Koronakis, V. & Hughes, C. Bacterial signal peptide-independent protein export: HlyB-directed secretion of hemolysin. *Sem. Cell Biol.* **4**, 7–16 (1993).
8. Gilson, L., Mahanty, H. K. & Kolter, R. Genetic analysis of an MDR-like export system: the secretion of colicin V. *EMBO J.* **9**, 3875–3894 (1990).
9. Nikaïdo, H. Multidrug efflux pumps of Gram-negative bacteria. *J. Bacteriol.* **178**, 5853–5859 (1996).
10. Zgurskaya, H. I. & Nikaïdo, H. Bypassing the periplasm: reconstitution of the actA/B multidrug efflux pump of *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **96**, 7190–7195 (1999).
11. Stone, B. J. & Miller, V. L. *Salmonella enteritidis* has a homolog of TolC that is required for virulence in BALB/C mice. *Mol. Microbiol.* **17**, 701–712 (1995).
12. Koronakis, V., Li, J., Koronakis, E. & Stauffer, K. Structure of TolC, the outer-membrane component of the bacterial type I efflux system, derived from two-dimensional crystals. *Mol. Microbiol.* **23**, 617–626 (1997).
13. Sharff, A. J., Koronakis, E., Luisi, B. & Koronakis, V. Oxidation of selenomethionine: Some MADness in the method! *Acta Crystallogr. D* **56**, 785–788 (2000).
14. Benz, R., Maier, E. & Gentshev, I. TolC of *Escherichia coli* functions as an outer membrane channel. *Zentralbl. Bakteriol.* **278**, 187–196 (1993).
15. Lupas, A. Predicting coiled-coil regions in proteins. *Curr. Opin. Struct. Biol.* **7**, 388–393 (1997).
16. Johnson, J. M. & Church, G. M. Alignment and structure prediction of divergent protein families: periplasmic and outer membrane proteins of bacterial efflux pumps. *J. Mol. Biol.* **287**, 695–715 (1999).
17. Pautsch, A. & Schulz, G. E. Structure of the outer membrane protein A transmembrane domain. *Nature Struct. Biol.* **5**, 1013–1017 (1998).
18. Schirmer, T. et al. Structural basis for sugar translocation through maltoporin channels at 3.1 Å resolution. *Science* **267**, 512–514 (1995).
19. Cowan, S. W. et al. Crystal structures explain functional properties of two *E. coli* porins. *Nature* **358**, 727–733 (1992).
20. Kreusch, A. et al. Structure of the membrane channel porin from *Rhodospseudomonas blastica* at 2.0 Å resolution. *Protein Sci.* **3**, 58–63 (1994).
21. Snijder, H. J. et al. Structural evidence for dimerization-regulated activation of an integral membrane phospholipase. *Nature* **401**, 717–721 (1999).
22. Buchanan, S. K. et al. Crystal structure of the outer membrane active transporter FepA from *Escherichia coli*. *Nature Struct. Biol.* **6**, 56–63 (1999).
23. Locher, K. P. et al. Transmembrane signalling across the ligand-gated FhuA receptor: crystal structures of free and ferrichrome-bound states reveal allosteric changes. *Cell* **95**, 771–778. (1998).
24. Ferguson, A. D. et al. Siderophore-mediated iron transport: Crystal structure of FhuA with bound lipopolysaccharide. *Science* **282**, 2215–2220 (1998).
25. Schulz, G. E. Porins: general to specific, native to engineered passive pores. *Curr. Opin. Struct. Biol.* **6**, 485–490 (1996).
26. Seo, J. & Cohen, C. Pitch diversity in α -helical coiled coils. *Proteins Struct. Funct. Genet.* **15**, 223–234 (1993).
27. Crick, F. H. C. The packing of α -helices: simple coiled coils. *Acta Crystallogr. B* **6**, 689–697 (1953).
28. Oas, T. G. & Endow, S. A. Springs and hinges: dynamic coiled coils and discontinuities. *Trends Biochem. Sci.* **19**, 51–54 (1994).
29. Unwin, N. Acetylcholine receptor channel imaged in the open state. *Nature* **373**, 37–43 (1995).
30. Schatz, G. & Dobberstein, B. Common principles of protein translocation across membranes. *Science* **271**, 1519–1526 (1996).
31. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
32. Terwilliger, T. C. & Berendzen, J. Automated structure solution for MIR and MAD. *Acta Crystallogr. D* **55**, 849–861 (1999).
33. de la Fortelle, E. & Bricogne, G. Maximum-likelihood heavy-atom parameter refinement in the MIR and MAD methods. *Methods Enzymol.* **276**, 472–494 (1997).
34. Collaborative Computer Project No. 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
35. Cowtan, K. in *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* Vol. 31 34–38 (Daresbury Laboratory, Daresbury, UK, 1994).
36. Brünger, A. T. et al. Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
37. Perrakis, A., Sixma, T. K., Wilson, K. S. & Lamzin, V. S. wARP: improvement and extension of crystallographic phases by weighted averaging of multiple refined dummy atomic models. *Acta Crystallogr. D* **53**, 448–455 (1997).
38. Bricogne, G. Direct phase determination by entropy maximization and likelihood ranking: status report and perspectives. *Acta Crystallogr. D* **49**, 37–60 (1993).
39. Carson, M. Ribbons. *Methods Enzymol.* **277**, 493–505 (1997).
40. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11**, 281–296 (1991).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank A. Thompson, K. Henderson, N. Duke, A. Joachimiak, Z. Otwinowski, W. Minor, K. Phillips, M. Symmons and F. von Delft for help during data collection and discussions; C. Calladine, M. Perutz, J. Li and C. Andersen for discussions; and G. Bricogne and P. Roversi for help with the BUSTER program. We acknowledge support for EU facility access for data collection at DESY (Hamburg) and ESRF (Grenoble). This work is supported by a Medical Research Council Programme grant (to V.K. and C.H.). A.S. and B.L. are supported by the Wellcome Trust.

Correspondence and requests for materials should be addressed to V.K. (e-mail: vk103@mole.bio.cam.ac.uk). Coordinates have been deposited in the RCSB protein data bank under accession code 1EK9.

Optical detection of meteoroidal impacts on the Moon

J. L. Ortiz*, P. V. Sada†, L. R. Bellot Rubio‡, F. J. Aceituno§, J. Aceituno||, P. J. Gutiérrez* & U. Thiele||

* Instituto de Astrofísica de Andalucía, CSIC, Aptdo 3004, 18080 Granada, Spain

† Departamento de Física y Matemáticas, Universidad de Monterrey, Av. Ignacio Morones Prieto 4500 Pte., San Pedro Garza García, Nuevo León, CP 66238, México

‡ Instituto de Astrofísica de Canarias, E-38200, La Laguna, Tenerife, Spain

§ C/Alhambra 40, 18004 Granada, Spain

|| Centro Astronómico Hispano Alemán, Aptdo 511, 04080 Almería, Spain

Impacts of meteoroids on the Moon should cause detectable optical flashes¹, but the population of objects that are big enough is very low, and hitherto no unambiguous impact flashes have been recorded. The flux of meteoroids associated with the Leonid meteor shower of 18 November 1999 was predicted to produce observable flashes on the night side of the Moon². Here we report the unambiguous detection of five such impact flashes, three of which were seen simultaneously by other observers³. We also observed a possible impact flash on 16 July 1999. All of the flashes were of very brief duration (<0.02 s), as expected for high-speed impacts.

A good knowledge of the sporadic meteoroid population and meteoroid streams is needed for a complete understanding of the Solar System and also for the safety of artificial satellites. Our current knowledge is primarily based on observational data from visual, photographic and radar meteor records, from spacecraft long-exposure facilities and from cratering rates on the Moon⁴. Studies of impacts of meteoroids on the night side of the Moon can complement the above techniques. The Moon can be regarded as a long-exposure facility with a huge collecting area, and therefore measurements of direct impacts may be very valuable in characterizing the current population of meteoroids in the vicinity of the Earth, especially those of high mass, whose impact probability in the small areas covered by the above-mentioned techniques is low. In addition, as the lunar atmosphere is extremely rarefied, useful information on the physics of hypervelocity impacts can be gathered, and properties of the impacting and surface material may be retrieved in the future, opening a new field for remote sensing of Solar System objects. Moreover, the nature and evolution of the sodium lunar atmosphere can be investigated by correlating these studies with sodium lunar emission measurements; reports on enhancements of the lunar sodium atmosphere as a result of the 1997 and 1998 Leonid meteor showers have been published recently^{5,6}.

In 1997, a search for impact events was begun and a first report was presented². Although the search failed to detect unambiguously any impact, it was pointed out that these studies would be very useful for the 1999 Leonids. Hence we requested telescope time at several observatories for the Leonid 1999 campaign and for preparation campaigns. Observing time was granted at the 0.8-m Schmidt telescope of Calar Alto observatory, the 1.5-m infrared telescope at Teide Observatory and the 0.9-m telescope at Sierra Nevada Observatory. We also set up several small telescopes for the Leonids. Bad weather and technical problems prevented us from observing with the larger telescopes. However, observations conducted by one of us (P.V.S.) from Monterrey (Mexico) were successful and five impacts were detected (Fig. 1), three of which were observed at times and lunar positions coincident with impacts videotaped by other observers³ (Table 1).

The Moon had a closest approach of 0.0002 AU (astronomical units) to the Leonid stream at 4:49 UT (universal time) on 18 November 1999 (ref. 3), near the midpoint of the observing



Figure 1 Images of the Leonid impact flashes. From Monterrey (Mexico) five impacts were recorded in about 90 minutes centred around 4:30 UT, on 18 November 1999. For those observations, a 0.20-m telescope equipped with a black and white charge-coupled device (CCD) video camera and a video recorder was used to monitor part of the night side of the Moon, away from the terminator. A radio set tuned at a short-wave station was used to record UTC signals simultaneously. The observations turned out to be difficult because the high background caused by the daylight side of the Moon did not allow the dark limb to be seen, making tracking corrections difficult and preventing a precise determination of impact location. After carefully reviewing the tape, five flashes were clearly seen (with signal-to-noise ratios ranging from 7 to 18), three of which were observed at times coincident with impacts videotaped by other observers³. The relevant data are summarized in Table 1. The flashes occurred mainly in half-frames (which correspond to integrations of 1/60 s). The magnitude estimates for them are 3, 4, 7, 6 and 5, based on the signals of recorded reference stars. Translated into energy per unit area reaching the telescope in the interval from 0.4 to 0.9 μm , this means 1.81×10^{-11} , 7.21×10^{-12} , 4.54×10^{-13} , 1.14×10^{-12} and $2.87 \times 10^{-12} \text{ J m}^{-2}$ with an estimated uncertainty of 30%. The half-frame afterglows in the two brightest impacts were very faint. The field of view of the camera was 8×6 arcmin, which corresponds to an average of $0.9 \pm 0.2 \times 10^6 \text{ km}^2$ of lunar area, the uncertainty arising from the inaccurate tracking. The images shown here are 53×53 arcsec enlargements of the half-frames around the impact sites (digitized using a frame-grabber). The telescope was somewhat defocused during the last two impacts.

period shown in Fig. 1. Such a distance to the three-revolution-old trail was equal to the Earth's 1966 distance to the two-revolution-old trail⁷. Therefore, the 1999 Moon's approach to the Leonids was very similar to that of the Earth in 1966 and we can expect that lunar infall fluxes in 1999 were of the same order as terrestrial fluxes in 1966. Therefore, the flashes recorded are indeed the result of Leonid meteoroids striking the Moon. The data presented here, or similar data, can be used to study the structure of the Leonid dust trails encountered by the Earth in 1999 (for example, the mass distribution in a mass range that is difficult to observe on Earth with conventional techniques) or some aspects of the physics of high-speed impacts, such as luminous efficiency (η).

A rough guess at the luminous efficiency can be derived from the data by assuming a power law for the distribution of masses in the Leonid stream and using the flux of meteoroids in 1966 (ref. 8) scaled down by a factor of four (as there are indications that the spatial density of the three-revolution-old trail must have decreased^{9,10} by a factor of two to five). Plausible indices of the mass distribution can range from $s = 1.8$ (obtained for the 1999 terrestrial meteors¹¹) to $s = 2.0$ (indicated for the 1966 storm^{8,12}). These values imply that the mass of a Leonid meteoroid whose impact rate in our field of view was one per 90 minutes corresponds to 9 kg and 1.3 kg, for $s = 1.8$ and $s = 2.0$ respectively. According to this, the brightest impact corresponded to a mass range of 9 kg to 1.3 kg and therefore, values in the order of 10^{-3} to 10^{-2} for the luminous efficiency would be indicated.

During our Leonid preparation campaign of July 1999, we recorded another signature which is possibly an impact flash. Figure 2 shows the candidate, detected by using a slightly different method and unrelated to the Leonids. It is possible that the bright spot is due to a cosmic ray hit or some kind of electronic noise. However, we took 5,000 images under a similar high background (by illuminating the dome) and we searched for areas that might look alike. We found three spots with one saturated pixel, but the number of pixels in the spots never exceeded 9, whereas our impact candidate occupies 14 pixels and is fairly round, as can be seen in the third panel of Fig. 2. A point spread function (PSF) fit to the suspected impact revealed that the shape of the function was similar to a usual PSF, although the wings of the intensity distribution were shorter than those of a typical long-exposure PSF. We believe that this difference stems from the fact that the impact flash was so brief that atmospheric turbulence could not generate the typical wings of long-integration PSFs. Indeed, the flashes reported for the Leonid impacts were very short, which strengthens our theory.

Specular reflection of sunlight from artificial satellites could cause very brief flashes. However, the lack of trail and the absence of another flash in the same frame or in other frames indicate that the feature is not likely to be due to a satellite, nor to space debris.

Assuming a luminous efficiency of order 10^{-3} , the mass required to generate the signal of the proposed July impact is 14 kg if we adopt the speed of sporadic meteoroids (20 km s^{-1}). According to the flux of large sporadic meteoroids¹³, an impact of a 14-kg body in the field of view should occur once every 200 hours of observations, which is approximately a factor of 10 more time than we have spent observing the Moon. This might indicate one or several of these

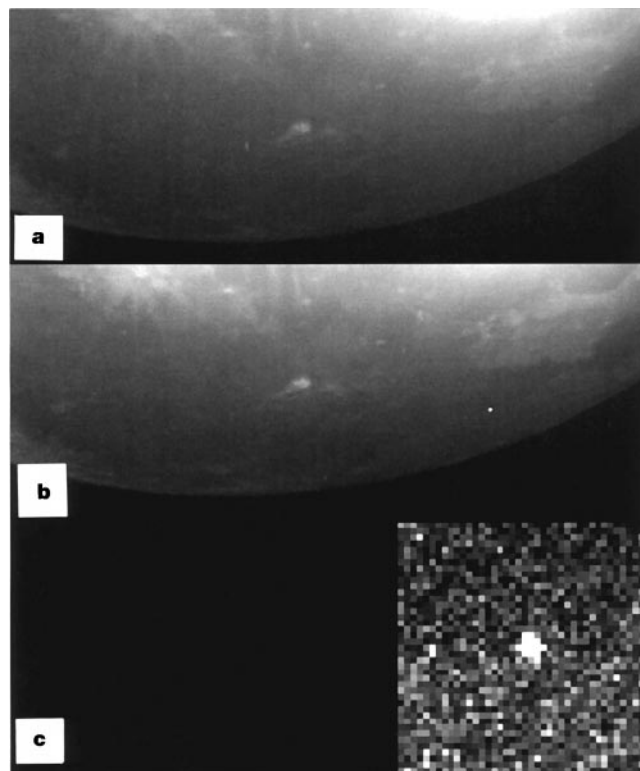


Figure 2 Images of a possible impact flash on 16 July 1999. To obtain these images we employed a thermoelectrically cooled CCD attached to the focal plane of the 0.8-m Schmidt telescope at Calar Alto Observatory. The CCD chip was a Kodak KAF-1600. We took sequences of 4-s exposures, similar to the procedure used in our previous work² to detect impacts. **a**, An exposure of the Moon's night side taken at 20:45:21 UT, a few seconds before a suspected impact took place. The Earth-lit Moon shows no sign of impact. The image scale of the CCD detector employed is 1.56 arcsec per pixel. The field of view is 13.2×5.3 arcmin. The Earth-lit Moon is seen here because the exposures were of considerable length. **b**, The image containing the suspected impact flash. The bright spot covers 14 pixels, two of which were saturated. The image was taken at 20:45:36 UT. **c**, An enlargement of the image resulting from the subtraction of image **a** from image **b**. White corresponds to points exceeding three times the standard deviation of the background. The detector counts from the event are somewhat uncertain, as there are two saturated pixels. Nevertheless, a PSF fit gives a flux of 169,400 counts, which means $4.5 \times 10^{-12} \text{ J m}^{-2}$ in the range from 0.4 to $0.9 \mu\text{m}$.

possibilities: (1) The luminous efficiency for that impact was somewhat higher than 10^{-3} and hence, the impactor's mass was considerably smaller than 14 kg. (2) The impactor had a higher speed than the average speed of sporadics and possibly belonged to any of the several slightly active meteoroid streams at that time (all of them have higher speeds than the sporadics). (3) Our chances of detecting such an impact were 10% but we were lucky enough to record the impact. (4) The sporadic flux magnitude¹³ is in error in

Table 1 Relevant impact data

Impact number	UTC (± 0.02 s)	Energy per unit area received on Earth (J m^{-2})	Location on the Moon	Other observers	UTC from other observers ³	Location on the Moon by other observers ³
1	03:49:40.38	1.81×10^{-11}	$5^\circ \text{N} \pm 5^\circ, 60^\circ \text{W} \pm 15^\circ$	Dunham, Palmer, Frankenberger	03:49:40.5 $\pm$ 0.4	$3^\circ \text{N} \pm 5^\circ, 48^\circ \text{W} \pm 5^\circ$
2	04:32:50.79	7.21×10^{-12}	$21^\circ \text{N} \pm 3^\circ, 51^\circ \text{W} \pm 3^\circ$			
3	04:34:49.52	4.54×10^{-13}	$21^\circ \text{N} \pm 3^\circ, 38^\circ \text{W} \pm 3^\circ$			
4	05:14:12.92	1.14×10^{-12}	$13^\circ \text{N} \pm 3^\circ, 57^\circ \text{W} \pm 4^\circ$	Dunham	05:14:12.93 $\pm$ 0.05	$15^\circ \text{N} \pm 2^\circ, 58^\circ \text{W} \pm 2^\circ$
5	05:15:20.22	2.87×10^{-12}	$20^\circ \text{N} \pm 5^\circ, 55^\circ \text{W} \pm 10^\circ$	Dunham	05:15:20.23 $\pm$ 0.05	$21^\circ \text{N} \pm 2^\circ, 59^\circ \text{W} \pm 2^\circ$

Summary of the impact data of Fig. 1 along with information from other observers³.

the 1 to 10 kg range. (5) The feature is unrelated to impacts. Of these options, we favour the possibility that the luminous efficiency was higher than 10^{-3} . Also, it is very likely that the impactor was a meteoroid belonging to any of the active streams rather than a sporadic.

Using the luminous efficiency that we derive (and even if it were in error by two orders of magnitude) we can rule out the re-proposed small-comet theory¹⁴ to explain the atmospheric holes in the Earth's day glow. Such minicometes would produce frequent and spectacular flashes when striking the Moon and we would have detected them. In order not to produce flashes, such small comets would have to be broken into very small pieces before reaching the lunar surface. □

Received 21 December 1999; accepted 5 May 2000.

1. Melosh, H. J. *et al.* Remote visual detection of impacts on the lunar surface. *Lunar Planet. Sci. Conf.* XXIV, 975–976 (1993).
2. Ortiz, J. L., Aceituno, F. J. & Aceituno, J. A search for meteoritic flashes on the Moon. *Astron. Astrophys.* 343, L57–L60 (1999).
3. Dunham, D. W. *IAU Circ.* No. 7320 (1999).
4. Grün, E., Zook, H. A., Fichtig, H. & Riese, R. H. Collisional balance of the meteoritic complex. *Icarus* 62, 244–272 (1985).
5. Huntent, D. M., Sprague, A. L. & Cremonese, G. Lunar Na during the 1997 Leonid meteor shower. *Bull. Am. Astron. Soc.* 30, 1115 (1998).
6. Verani, S., Barbieri, C., Benn, C. & Cremonese, G. Possible detection of meteor stream effects on the lunar sodium. *Planet. Space Sci.* 46, 1003–1006 (1998).
7. Asher, D. J. Leonid meteor storms of 1833 and 1966. *Mon. Not. R. Astron. Soc.* 307, 919–924 (1999).
8. Brown, P., Šimek, M. & Jones, J. Radar observations of the Leonids: 1964–1995. *Astron. Astrophys.* 322, 687–695 (1997).
9. McNaught, R. H. & Asher, D. J. Leonid dust trails and meteor storms. *J. Int. Meteor. Organizat.* 27, 85–102 (1999).
10. Brown, P. Evolution of two periodic meteoroid streams: the Perseids and Leonids. Thesis, Univ. Western Ontario (1999).
11. Arlt, R., Bellot Rubio, L. R., Brown, P. & Gyssens, M. First global analysis of the 1999 Leonid storm. *J. Int. Meteor. Organizat.* 27, 286–295 (1999).
12. Brown, P. The Leonid meteor shower: Historical visual observations. *Icarus* 138, 287–308 (1999).
13. Cepelch, Z. Impacts of meteoroids larger than 1 m into the Earth's atmosphere. *Astron. Astrophys.* 286, 967–970 (1994).
14. Frank, L. A. & Sigwarth, J. B. Trails of OH emissions from small comets near Earth. *Geophys. Res. Lett.* 24, 2435–2438 (1997).

Acknowledgements

We are grateful to V. Costa, J. Gutiérrez, H. Boehnhardt, D. Hamilton, T. Herbst, J. J. Lopez-Moreno, A. J. Castro-Tirado, E. García-Lobo, F. Prada, F. Hoyos, A. Porcel, A. García and the amateur astronomers at Observatorio de Mallorca, for their help and support. This work was partially supported by project ESP97-1773-CO3-01.

Correspondence and requests for materials should be addressed to J.L.O. (e-mail: ortiz@iaa.es).

Hidden symmetries in the energy levels of excitonic 'artificial atoms'

M. Bayer*, O. Stern*, P. Hawrylak*, S. Fafard† & A. Forchel*

* Technische Physik, Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany

† Institute for Microstructural Science, National Research Council of Canada, Ottawa, K1A 0R6, Canada

Quantum dots^{1–7} or 'artificial atoms' are of fundamental and technological interest—for example, quantum dots^{8,9} may form the basis of new generations of lasers. The emission in quantum-dot lasers originates from the recombination of excitonic complexes, so it is important to understand the dot's internal electronic structure (and of fundamental interest to compare this to real atomic structure). Here we investigate artificial electronic structure by injecting optically a controlled number of electrons and holes into an isolated single quantum dot. The

charge carriers form complexes that are artificial analogues of hydrogen, helium, lithium, beryllium, boron and carbon excitonic atoms. We observe that electrons and holes occupy the confined electronic shells in characteristic numbers according to the Pauli exclusion principle. In each degenerate shell, collective condensation of the electrons and holes into coherent many-exciton ground states takes place; this phenomenon results from hidden symmetries (the analogue of Hund's rules for real atoms) in the energy function that describes the multi-particle system. Breaking of the hidden symmetries leads to unusual quantum interferences in emission involving excited states.

The first step in constructing excitonic artificial atoms is to understand the confined states of electrons and holes in the self-assembled quantum dots (see Fig. 1a). This understanding can be obtained by high-excitation spectroscopy¹⁰ applied to samples containing millions of dots. In highly excited dots, several electron and hole energy levels are occupied. These levels may be degenerate due to geometrical and dynamical symmetries and therefore form electronic shells $s, p, \dots$ (ref. 6). The shells are populated with carriers according to the Pauli principle. In the spectra, distinct emission

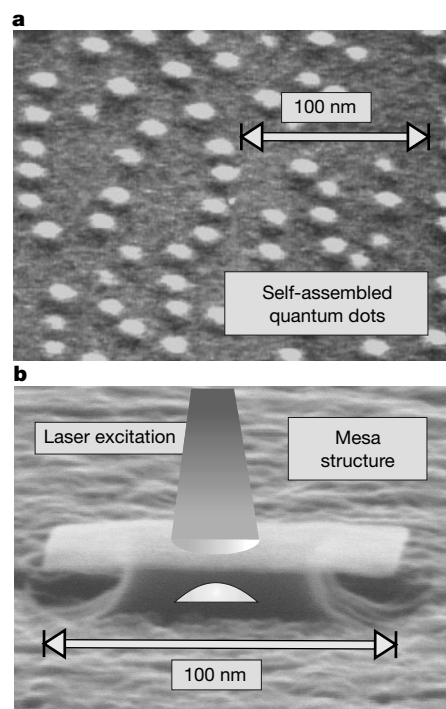


Figure 1 Scanning electron micrographs illustrating the experimental technique used for studying single self-assembled quantum dots. **a**, Scanning electron micrograph of a GaAs semiconductor layer on which $\text{In}_{0.60}\text{Ga}_{0.40}\text{As}$ self-assembled quantum dots with a density of about 10^{10} cm^{-2} have been grown by molecular beam epitaxy. To permit their microscopic observation these dots—unlike those used for spectroscopy—have not been covered by a GaAs cap layer. To a good approximation, all quantum dots have the same shape exhibiting rotational symmetry. However, their size varies by a few nanometres around an average diameter of 15 nm. This inhomogeneity results in a considerable broadening of the emission lines in spectroscopic studies. **b**, To avoid this broadening we have studied the emission of a single quantum dot. Lithographic techniques were used to fabricate small mesa structures on samples capped by a GaAs layer. The lateral mesa size was reduced to such an extent ($<100 \text{ nm}$) that only a single dot is contained in it. These mesa structures have been studied by photoluminescence spectroscopy at low temperature. A laser beam (shown schematically as a truncated cone above the mesa) injects a controlled number of electrons and holes into the dot indicated by the lens shape, and the emission spectrum of this complex is recorded. To reduce sample heating under optical excitation, the structures are held in superfluid helium at about 1.2 K. After dispersion by a monochromator, the emission is detected by a CCD (charge-coupled device) camera.

features appear which correspond to recombination of excitons from several confined shells.

Typical spectra of an ensemble of $\text{In}_{0.60}\text{Ga}_{0.40}\text{As}$ self-assembled quantum dots are shown in Fig. 2. At low excitation powers P_{ex} , emission from the ground shell as well as from the two-dimensional wetting layer is observed. The s shell is two-fold spin-degenerate, and can be occupied by not more than two excitons. Increasing P_{ex} results in complete s -shell filling and saturates its emission. Simultaneously, emission from the p shell appears. However, in these spectra the effects of Coulomb interactions are blurred by inhomogeneous broadening. Optical spectroscopy of single quantum dots^{11–23} has been successfully applied to suppress inhomogeneous broadening and to resolve correlation effects. We use single-dot spectroscopy—as sketched in Fig. 1b—to study changes of the emission spectra when the number of excitons is varied in a controlled manner.

Figure 3 shows the emission intensity from the s and p shells of a single dot as a function of excitation power and energy. The low-excitation spectra exhibit a single sharp emission line, X, which can be attributed to recombination of a single electron–hole pair in the s shell. With increasing illumination, biexciton emission X_2 becomes dominant at the low-energy side of X while the exciton line fades away. Then the biexciton emission decreases, while a new s -shell emission line emerges at slightly lower energies. Simultaneously, intense emission from the p shell appears at an energy 50 meV higher.

In the further evolution of the spectra with excitation power, two characteristic features are to be noted. (1) In the p -shell emission, only a very restricted number of spectral lines are observed. First, a single emission line appears (X_3). This emission line is replaced by two lines (X_4), followed by two other lines with very similar energies (X_5). Finally, these two lines are replaced by a single line (X_6). The energies of all these lines are approximately the same, and the splittings between them are small as compared to the splitting between the quantum-dot shells. (2) In contrast, there are dramatic changes in the number of lines of the s -shell emission and in their energies. At low p -shell population a single s -shell emission line (X_3) is observed. This line is replaced by a broad band (X_4) consisting of four strong emission lines, when the emission from the p shell splits into two. On increasing the illumination, we come to a situation (X_5) in which the s -shell emission almost completely vanishes. Increasing the excitation further leads to a single emission line (X_6) from the s shell when there is also a single line from the p shell.

To facilitate the comparison with theory, we have selected six spectra from Fig. 3 (indicated by red lines), in which a certain set of

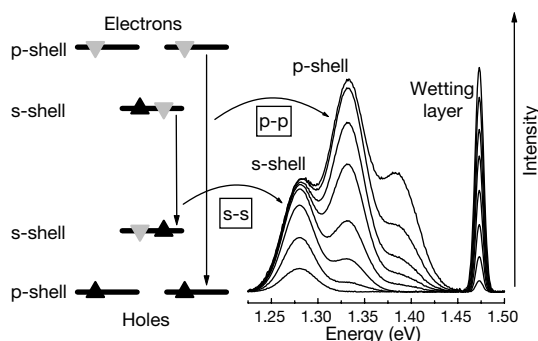


Figure 2 State filling spectroscopy on quantum dots. On the left is a scheme of the dot energy levels, their occupation by carriers and the radiative transitions. Spin orientations of electrons and holes: grey triangles, spin-down; black triangles, spin-up. On the right are typical emission spectra resulting from these transitions for an ensemble of $\text{In}_{0.60}\text{Ga}_{0.40}\text{As}$ quantum dots; these spectra were recorded at different excitation powers (an Ar-ion laser was used).

emission lines has fully evolved, while the set observed for lower (higher) excitation has disappeared (not yet appeared). These spectra shown in Fig. 4 correspond most closely to quasi-equilibrium conditions in which the mean number of excitons in the dot is an integer. In the following, they are compared with calculated spectra for different exciton occupations. These calculations reveal the fundamental principle of hidden symmetries as a rule determining the electronic structure of electron–hole complexes in quantum dots. In this sense, the hidden symmetries are the analogue of Hund's rules in atomic physics. The hamiltonian of the interacting few particle system is given by^{24,25}

$$H = \sum_i E_i^e c_i^\dagger c_i + \sum_i E_i^h d_i^\dagger d_i - \sum_{ijkl} \langle ij | V_{eh} | kl \rangle c_i^\dagger d_j^\dagger d_k c_l + \frac{1}{2} \sum_{ijkl} \langle ij | V_{ee} | kl \rangle c_i^\dagger c_j^\dagger c_k c_l + \frac{1}{2} \sum_{ijkl} \langle ij | V_{hh} | kl \rangle d_i^\dagger d_j^\dagger d_k d_l$$

where $c_i^\dagger$ and $d_i^\dagger$ (c_i and d_i) are the creation (annihilation) operators for electrons and holes. $E_i^{e/h}$ are the electron/hole single particle energies and V_{mn} , $m, n = e, h$ are the interparticle Coulomb interactions.

The interband optical processes are described by the polarization operator $P^+ = \sum_i c_i^\dagger d_i^\dagger$, where P^+ annihilates a photon and creates an electron–hole pair. The main question arises when populating degenerate shells: in atoms, Hund's rules determine which few-electron configurations give dominant contribution to the ground state. Here we are populating shells with two types of particles, electrons and holes. The ground state of this bi-polar system can be found by examining the dynamics of the polarization operator, that is, the commutator of P^+ with the hamiltonian. For degenerate shells and symmetric interactions, $V_{ee} = V_{hh} = -V_{eh}$, the commutator has a simple form²⁵:

$$[H, P^+] = +E_X P^+$$

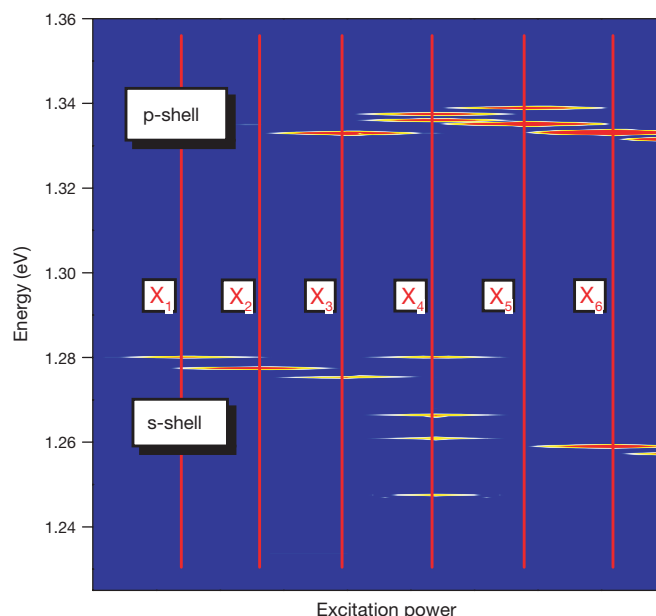


Figure 3 Contour plot of the variation of the emission of an $\text{In}_{0.60}\text{Ga}_{0.40}\text{As}$ single quantum dot with excitation power and with energy. Bright regions indicate strong emission intensities, blue regions low intensities. When optically exciting far above the bandgap, carrier relaxation involving multiple phonon emission processes leads to considerable sample heating, which causes the system to be in strong non-equilibrium. To reduce heating, a Ti-sapphire laser was used as excitation source. Its energy was tuned to $E = 1.470$ eV, corresponding to emission close to the bottom of the wetting layer (see Fig. 2). The excitation power P_{ex} was varied between 50 nW and 5 mW.

Here E_X is the energy of an exciton in a given shell. This commutation relation implies a symmetry in the system which is not obvious and is therefore called 'hidden symmetry'. Owing to the commutator, coherent multiplicative states $|N\rangle = (P^\dagger)^N |0\rangle$ of N electron-hole pairs are exact ground eigenstates of the hamiltonian. The important manifestation of hidden symmetry is the linear dependence of the multiplicative state energy on the number of excitons: $E_N = N \cdot E_X$. Hence the energy to add or subtract an exciton to a shell does not depend on the filling of this shell.

But how does hidden symmetry affect the emission spectra of excitonic artificial atoms? The properties of the spectra are determined by the initial and the final states of the electron-hole recombination. As an example we consider the radiative decay of a four- to a three-exciton state. Two of the excitons occupy the s shell, the other two occupy the p shell. There are two principal radiative decay channels: either an exciton from the s shell or an exciton from the p shell can recombine. For p -shell recombination both initial and final states are multiexciton ground states, for which the hidden symmetries are valid. Therefore the emission spectrum should not depend on the filling of the p shell. In contrast, for s -shell recombination the hidden symmetry applies only for the initial state, whereas the final state is not a multiplicative state due to the vacancy in the s shell. Therefore strong variations of the s -shell recombination with exciton occupation are expected.

We have calculated the emission spectra for up to $N = 6$ excitons in a dot corresponding to full occupation of the s and p shells. In modelling the multiexciton states Coulomb interactions are treated exactly for excitons within a shell while scattering to higher shells is taken into account perturbatively. From these N exciton states the emission spectra $L_N(\omega)$ are calculated by Fermi's golden rule:

$$L_N(\omega) = \sum_f |\langle N-1, f | P^- | i, N \rangle|^2 \cdot \delta(E_N^i - E_{N-1}^f - \hbar\omega)$$

where we assume a quasi-equilibrium for the electron-hole system. E_i^N and E_f^{N-1} are the energies of the initial (i) (final (f)) N ($N-1$)

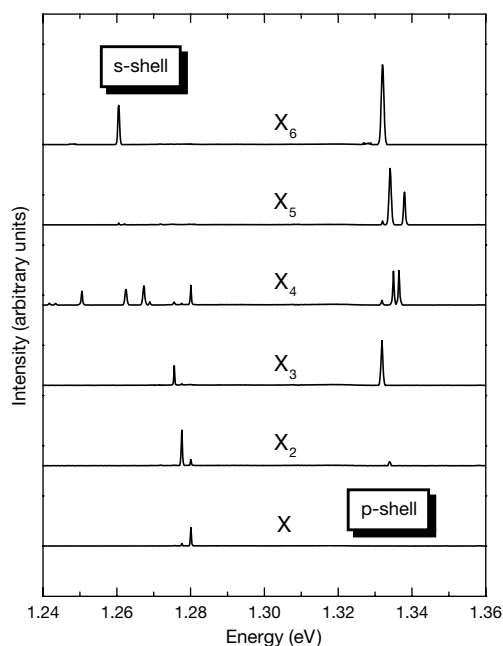


Figure 4 Photoluminescence spectra of an $\text{In}_{0.60}\text{Ga}_{0.40}\text{As}$ single quantum dot for varying excitation powers. The spectra have been selected from Fig. 2 to approximate most closely the situations in which the time-averaged dot occupation is given by an integer number of excitons. The emission powers used for recording the spectra were (from bottom to top): 200 nW, 800 nW, 2.9 μW , 17 μW , 310 μW , 4.3 mW.

exciton states. The results of these calculations for parameters typical of self-assembled quantum dots are shown in Fig. 5.

The two lowest traces show the emission spectra when only the s shell is occupied; the other traces give the emission for different p -shell occupations (from one up to four excitons) with the s shell fully occupied. Owing to the hidden symmetries the p -shell emission indeed varies only slightly with the shell occupation. The splitting of the emission line for the 4 (5) exciton complexes arises from different spin configurations: In the 4X (5X) complex the two electrons (holes) in the p shell form either triplet-triplet or singlet-singlet configurations in the initial (final) state of the recombination. These configurations are almost degenerate and the mean energy of the p -shell emission does not depend on its population. Therefore it also gives no clear signatures for the exciton number in the dot. In contrast, the s -shell recombination does give such signatures, which for dot occupations by more than two excitons can be summarized as follows.

(1) The three-exciton complex (excitonic lithium) recombines primarily into spin-polarized final states with one electron (hole) in the s orbital and one in the p orbital. Its recombination energy is close to the 1X recombination energy because an exciton involving an electron (and hole) with spin opposite to the spin of the other electrons (and holes) is removed.

(2) For four excitons (excitonic beryllium), that is, a half-filled p shell, recombination takes place to a band of states corresponding to excited states of the three-exciton complex. The majority of these states involves a half-filled spin-polarized p shell and an s state with an opposite spin electron (or hole). One can easily generate states in the electron-hole system by merely flipping spin. The band of these states becomes optically active by mixing with one optically active configuration due to Coulomb interaction. The final result shows four intense emission lines from the s shell.

(3) For five excitons (excitonic boron) the most counter-intuitive result is obtained: the s -shell emission is suppressed due to destructive interference effects in its transition matrix element.

(4) For six excitons in the quantum dot (excitonic carbon), the s -

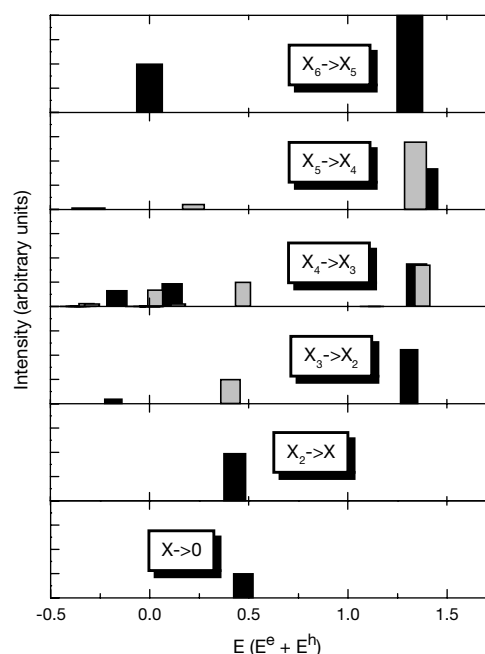


Figure 5 Calculated emission spectra of a quantum dot for different exciton occupations N indicated at each trace. The energy, E , is given in units of the kinetic quantization energy of electron and hole, $(E^e + E^h)$. The heights of the columns give the intensities of the emission lines. Different column shadings have been used to distinguish the transitions involving either singlet-singlet (black) or triplet-triplet (grey) spin configurations.

shell emission is recovered with an energy significantly lowered in comparison with the emission line of a single exciton due to exchange interaction (bandgap renormalization).

The calculations are in good agreement with the experimental observations. As the emission in quantum-dot lasers originates from recombination of excitonic complexes, these calculations should be applicable to the electronic structure of such lasers. The demonstration of 'atomic spectroscopy' in a solid-state system might also open up possibilities of using quantum dots for quantum information processing²⁶. Quantum dots offer flexibility in designing and controlling their interaction with optical pulses. Optically driven coupled quantum dots might lead to the realization of quantum gates, a building block of a quantum computer. A major obstacle to realizing an optically driven quantum computer in semiconductors is the fast dephasing of interband polarization due to carrier-carrier scattering. The hidden symmetries in excitonic artificial atoms imply that the four-particle correlations in the hamiltonian—which cause the dephasing—cancel. Therefore the realization of hidden symmetries may prove useful for the construction of decoherence-free (as far as carrier-carrier interactions are concerned) quantum dynamics. □

Received 20 January; accepted 28 April 2000.

- Jacak, L., Hawrylak, P. & Wojs, A. *Quantum Dots* (Springer, Berlin, 1998).
- Grundmann, M., Bimberg, D. & Ledentsov, N. N. *Quantum Dot Heterostructures* (Wiley & Sons, New York, 1998).
- Kastner, M. A. Artificial atoms. *Phys. Today* **46**, 24–31 (1993).
- Ashoori, R. C. Electrons in artificial atoms. *Nature* **379**, 413–419 (1996).
- Alivastos, A. P. Semiconductor clusters, nanocrystals, and quantum dots. *Science* **271**, 933–937 (1996).
- Banin, U., Cao, Y., Katz, D. & Millo, O. Identification of atomic-like electronic states in indium arsenide nanocrystal quantum dots. *Nature* **400**, 542–544 (1999).
- Fafard, S. *et al.* Red-emitting semiconductor quantum dot lasers. *Science* **274**, 1350–1353 (1996).
- Eaglesham, D. J. & Cerullo, M. Dislocation-free Stranski-Krastanow growth of Ge on Si(100). *Phys. Rev. Lett.* **64**, 1943–1946 (1990).
- Leonard, D. K., Pond, K. & Petroff, P. M. Critical layer thickness for self-assembled InAs islands on GaAs. *Phys. Rev. B* **50**, 11687–11692 (1994).
- Raymond, S. *et al.* Exciton droplets in zero dimensional systems in a magnetic field. *Solid State Commun.* **101**, 883–889 (1997).
- Marzin, J.-Y. *et al.* Photoluminescence of single InAs quantum dots obtained by self-organized growth on GaAs. *Phys. Rev. Lett.* **73**, 716–719 (1994).
- Zrenner, A. *et al.* Quantum dots formed by interface fluctuations in AlAs/GaAs coupled quantum well structures. *Phys. Rev. Lett.* **72**, 3382–3385 (1994).
- Brunner, K. *et al.* Sharp-line photoluminescence and two-photon absorption of zero-dimensional biexcitons in a GaAs/AlGaAs structure. *Phys. Rev. Lett.* **73**, 1138–1141 (1994).
- Grundmann, M. *et al.* Ultrasharp luminescence lines from single quantum dots. *Phys. Rev. Lett.* **74**, 4043–4046 (1995).
- Nirmal, M. *et al.* Fluorescence intermittency in single cadmium selenide nanocrystals. *Nature* **383**, 802–804 (1996).
- Gammon, D. *et al.* Homogeneous linewidths in the optical spectrum of a single gallium arsenide quantum dot. *Science* **273**, 87–90 (1996).
- Gammon, D. *et al.* Nuclear spectroscopy in single quantum dots: nanoscopic Raman scattering and nuclear magnetic resonance. *Science* **277**, 85–88 (1997).
- Bayer, M. *et al.* Exciton complexes in In_xGa_{1-x}As/GaAs quantum dots. *Phys. Rev. B* **58**, 4740–4753 (1998).
- Landin, L. *et al.* Optical studies of individual InAs quantum dots in GaAs: few-particle effects. *Science* **280**, 262–264 (1998).
- Dekel, E. *et al.* Multiexciton spectroscopy of a single self-assembled quantum dot. *Phys. Rev. Lett.* **80**, 4991–4994 (1998).
- Bayer, M. *et al.* Electron and hole g factors and exchange interaction from studies of the exciton fine structure in In_{0.60}Ga_{0.40}As quantum dots. *Phys. Rev. Lett.* **82**, 1748–1751 (1999).
- Bonadeo, N. H. *et al.* Coherent optical control of the quantum state of a single quantum dot. *Science* **282**, 1473–1476 (1998).
- Toda, Y. *et al.* Efficient carrier relaxation mechanism in InGaAs/GaAs self-assembled quantum dots based on the existence of continuum states. *Phys. Rev. Lett.* **82**, 4114–4117 (1999).
- Hawrylak, P. Excitonic artificial atoms: Engineering optical properties of quantum dots. *Phys. Rev. B* **60**, 5597–5608 (1999).
- Wojs, A. & Hawrylak, P. Exciton-exciton interaction in highly excited self-assembled quantum dots. *Solid State Commun.* **100**, 487–491 (1996).
- Hawrylak, P., Fafard, S. & Wasilewski, Z. Engineering quantum states in self-assembled quantum dots for quantum information processing. *Condensed Matter News* **7**, 16–23 (1999).

Acknowledgements

This work was supported by the State of Bavaria. P.H. thanks the Alexander von Humboldt Stiftung for support.

Correspondence and requests for material should be addressed to M.B. (e-mail: mbayer@physik.uni-wuerzburg.de).

Optical emission from a charge-tunable quantum ring

R. J. Warburton^{*,†}, C. Schäfflein^{*}, D. Haft^{*}, F. Bickel^{*}, A. Lorke^{*}, K. Karrai^{*}, J. M. Garcia[‡], W. Schoenfeld[‡] & P. M. Petroff

^{*} Center for NanoScience and Sektion Physik, Ludwig-Maximilians-Universität, Geschwister-Scholl-Platz 1, 80539 München, Germany

[‡] Materials Department and QUEST, University of California, Santa Barbara, California 93106, USA

[†] Present address: Department of Physics, Heriot-Watt University, Edinburgh EH14 4AS, UK

Quantum dots or rings are artificial nanometre-sized clusters that confine electrons in all three directions. They can be fabricated in a semiconductor system by embedding an island of low-bandgap material in a sea of material with a higher bandgap. Quantum dots are often referred to as artificial atoms because, when filled sequentially with electrons, the charging energies are pronounced for particular electron numbers^{1–3}; this is analogous to Hund's rules in atomic physics. But semiconductors also have a valence band with strong optical transitions to the conduction band. These transitions are the basis for the application of quantum dots as laser emitters⁴, storage devices^{5–7} and fluorescence markers⁸. Here we report how the optical emission (photoluminescence) of a single quantum ring changes as electrons are added one-by-one. We find that the emission energy changes abruptly whenever an electron is added to the artificial atom, and that the sizes of the jumps reveal a shell structure.

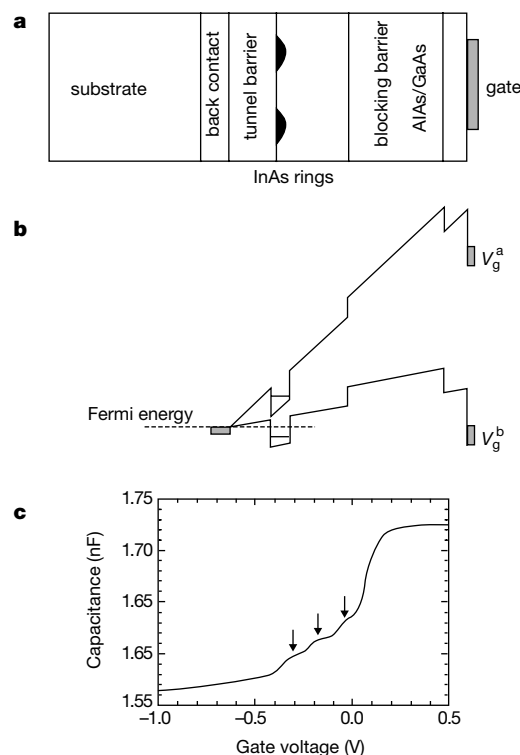


Figure 1 The semiconductor device for charging quantum rings with electrons. **a**, The layer structure. The tunnel barrier is 25 nm thick, and the separation between back contact and surface is 175 nm, so that a change in gate voltage ΔV_g implies a change in electrostatic potential (in eV) of $\Delta V_g/7$. **b**, The band diagram at two different voltages, V_g^a and V_g^b . **c**, The capacitance of the device used for the optical experiments (gate area 1.8 mm²) at 4.2 K. The arrows mark three charging peaks.

The nanometre-sized clusters for our experiments were made by self-assembly. We initially grew InAs quantum dots on GaAs by exploiting the strain-driven change from two-dimensional to three-dimensional growth which occurs after a coverage of 1.5 monolayers of InAs. A pause was introduced in the overgrowth with GaAs during which In migrated, causing the dots to become ring-like^{9,10}. The rings have an *s*-like ground state and a *p*-like excited state, as for dots, so that the general features of shell filling are the same for both dots and rings. The Aharonov–Bohm effect, an interference of electron waves in a quantum ring, is not relevant here as we do not apply a magnetic field to generate a phase difference between clockwise and anticlockwise paths around the ring.

We were faced with two experimental challenges. The first was to design a structure to load quantum rings sequentially with electrons. Our concept is to embed the InAs rings between highly doped GaAs and a surface gate electrode¹¹, as shown in Fig. 1a. At large, negative gate voltages (V_g^a in Fig. 1b), the ring level lies well above the Fermi energy and is unoccupied. At a more positive V_g , the ring level is resonant with the Fermi energy and electrons tunnel into and out of the ring. A further increase in V_g (V_g^b in Fig. 1b) traps one electron in the ring. We can monitor the tunnelling by using the capacitance between gate and back contact^{12,13}. Three charging peaks can be made out in the capacitance of a large ensemble of rings (Fig. 1c). The rise in capacitance at $V_g = 0.1$ V corresponds to tunnelling into the thin InAs layer between the rings (the wetting layer).

The second challenge was to measure a single ring. This is important as otherwise the 20-meV inhomogeneous broadening in the photoluminescence (PL) obscures the shifts expected on charging. We measured PL with a low-temperature confocal microscope which has a diffraction-limited spatial resolution of 610 nm at a wavelength of 950 nm. Our ring densities are $\sim 5 \times 10^9 \text{ cm}^{-2}$, implying that several tens of rings lie in the focus. Individual rings can be selected through their emission energy: we looked for rings in the low-energy tail of the spectral distribution. In order to investigate rings closer to the peak of the ensemble distribution, we increased the spatial resolution by defining 300-nm-sized holes in an aluminium film on the sample surface. Even in this case, we always detected emission from several rings. However, as each ring has unique charging voltages, the V_g dependence of the PL allows us to determine which PL lines belong to which ring. We present here data from a sample without a metal mask as the signal/noise ratio is superior.

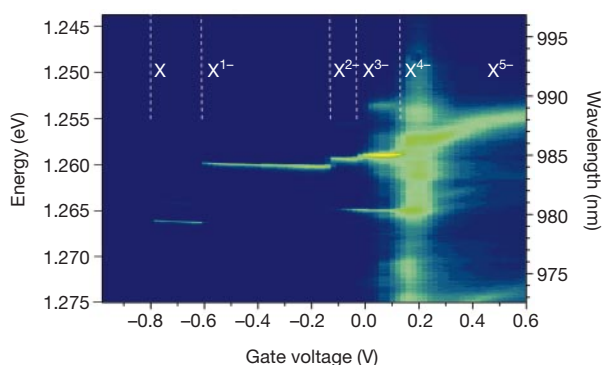


Figure 2 A colour-scale plot of the photoluminescence (PL) versus gate voltage at 4.2 K. Here dark blue, green and yellow correspond to low, medium and high signals, respectively. PL was excited by generating electron–hole pairs in the wetting layer with a laser diode emitting at 822 nm wavelength. The pump power was kept below 1 μW in order to populate the rings with no more than one exciton at a given time. The PL was dispersed with a 0.25-m grating spectrometer and detected with a cooled CCD camera, a set-up with spectral resolution 0.2 nm. The main features are from a single ring. The PL at 1.265 eV (980 nm), however, pronounced around 0.0 V, comes from a second ring.

We recorded the PL as a function of V_g . The data are represented as a colour-scale plot in Fig. 2. At $V_g \sim -0.7$ V there is a single, sharp peak which is the emission from a single ring. At -0.6 V the PL energy decreases abruptly. The jump in energy occurs because an additional electron becomes trapped in the ring. This voltage, -0.6 V, is more negative than the first charging voltage of the ensemble average (Fig. 1c), and this correlates with the particularly low emission energy of the ring in Fig. 2. At higher V_g , further steps can be made out in the PL. In each case, a jump arises when an additional electron is added to the ring. The first jump at -0.6 V, 6.0 meV in energy, represents the binding energy of a singly charged exciton (X^{1-}). The second jump represents the energy needed to add an additional electron to the X^{1-} to form X^{2-} , and so on. The first jump in energy is large, the second small, the third also small, and the fourth reasonably large. This is an optical manifestation of Hund's rules for electron charging. Whenever an electron can be added with the benefit of exchange energy, as is the case for the second and third steps, the charging energy is small. Whenever an electron is added to complete a sub-shell, the charging energy is large.

On moving from X^{1-} to X^{2-} , a satellite appears on the low-energy side of the main peak. The satellite can be just made out in the colour-scale plot, and is a clear feature in the individual spectra, a few of which are shown in Fig. 3. We can be confident that the two emission lines come from the same ring because the satellite

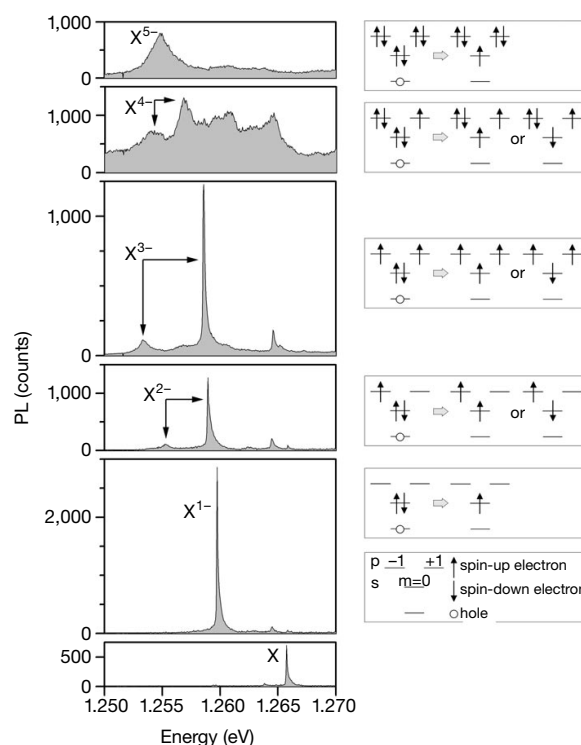


Figure 3 Photoluminescence from a single quantum ring for different charge states. Spectra are shown at $V_g = -0.76, -0.16, -0.10, 0.40, 0.22$ and 0.50 V, corresponding to emission from the $X, X^{1-}, X^{2-}, X^{3-}, X^{4-}$ and X^{5-} excitons, respectively. The PL red-shifts on charging, and a satellite occurs on the low-energy side of the main peak for X^{2-} and X^{3-} . Note that the other, weaker emissions (for instance on the high-energy side of the main X^{2-} PL) belong to another ring, and that the asymmetry in the sharp lines is an artefact of the spectrometer. The broad peak between the X^{3-} main peak and the satellite may be emission from an excited initial state which recombines before it can relax. The level diagrams show the emission processes from initial state to possible final states. We consider *s* and *p* states in the conduction band with angular momenta of 0 and ± 1 , and the *s* state in the valence band. In the experiment, both electrons and holes are unpolarized. To avoid redundancy in the figure, we sketch cases with electrons preferentially polarized up and an unpolarized hole.

emerges exactly when the PL changes from X^{1-} to X^{2-} (see also Fig. 4). Also, the main line weakens at the transition but the total intensity stays constant to within 10%. We claim that the X^{2-} has two lines because there are two different final states¹⁴, as shown in the level diagrams in Fig. 3. For X^{2-} , the two possible final states have either parallel spins (a triplet state) or antiparallel spins (a singlet state). In exact analogy to the excited states of the helium atom, these two states are separated by twice the exchange energy, $2X_{sp}$, and have degeneracies of 3 and 1, respectively. We measure the splitting to be 3.6 meV, that is $X_{sp} = 1.8$ meV. The intensity ratio is $7 \pm 2 : 1$, not the $3 : 1$ expected from the degeneracies. This could be due to the electron–hole exchange interaction which tends to align the spins in the initial state, favouring the triplet final state^{15,16}.

For X^{3-} a similar argument applies; in this case the final state is analogous to the excited lithium atom. We have diagonalized the interaction matrix for three spin-1/2 electrons, and calculate a splitting of $3X_{sp}$ and an intensity ratio $2 : 1$. Experimentally, the splitting is 1.44 times larger than for the X^{2-} , remarkably close to the predicted 1.5, and the satellite becomes more intense, 2 ± 0.5 times weaker than the main peak. It is surprising that a treatment of the Coulomb interactions with perturbation theory works so well. This approach is justified when the electron and hole wavefunctions extend less than the excitonic Bohr radius, 10 nm; yet here the circumference of the ring is ~ 100 nm. The X^{3-} is the most complicated case; X^{4-} is one electron short of a filled p shell and should behave like X^{2-} . In the data, the splitting for X^{4-} returns almost to the value for X^{2-} but the PL becomes very broad.

At large and negative V_g , the PL disappears; we interpret this as field ionization of the excitons. The electron tunnelling time, τ_t , decreases rapidly with increasing $|V_g|$, eventually becoming smaller than the recombination time, τ_r . At a smaller $|V_g|$ ($\tau_t > \tau_r$), the 0-electron and the 1-electron states become degenerate and electron tunnelling occurs. We detect this as a jump in the PL energy. Figure 4 shows how the high-energy peak weakens and the low-energy peak strengthens with small steps in V_g . It is puzzling that we observe both X and X^{1-} over a large range of V_g , some 20 mV, corresponding to a change in electrostatic energy of about 3 meV. This energy is much larger than the thermal energy (0.36 meV). Heating from the laser is unlikely as an explanation, as broad single electron tunnelling peaks have been observed in transport experiments on similar samples¹⁷. The 3-meV energy scale is also much larger than ring–ring interactions (1 meV) which are significantly screened by the adjacent back contact¹³. To account for this result, we propose that there are temporal fluctuations in the potential (on a timescale large compared to τ_t), caused by some sort of impurity state close to the ring. The effect is analogous to the spectral diffusion observed in the PL from single CdSe dots¹⁸. Figure 4 shows also the transition from X^{1-} to X^{2-} , and this too shows a gradual cross-over from one line to the other. Given the 3-meV fluctuations, it is at first sight surprising that we see sharp PL lines at all. The explanation is that the PL energy is very sensitive to the charge in the ring, but only weakly

sensitive to the electric field (almost flat plateaux in Fig. 2). The combination of the fluctuations with the slight V_g -dependence of the PL energy constitutes a broadening mechanism. We predict this to be about 20 μ eV for our sample, comparable in fact to the expected homogeneous linewidth¹⁹.

For highly charged excitons, X^{2-} and above, the final state after emission of a photon is an excited state. In the simplest case, the ring emits a photon and then some time later relaxes into the ground state. But if the final state relaxes very quickly, the emission will be broadened by the energetic uncertainty of the final state. Therefore, we can use the PL linewidth as a measure of the relaxation rates of the final states. For X^{2-} , the satellite is clearly broader (full-width at half-maximum, FWHM, 0.6 meV) than the main line (FWHM < 0.25 meV), implying a singlet decay time of 1.1 ps and that the singlet decays faster than the triplet. The different decay rates are a consequence of the spin: the triplet state has to flip a spin to relax, the singlet state does not. For the singlet, phonon scattering—which preserves electron spin—is the likely mechanism for the relaxation. Relaxation by phonon scattering is therefore fast on the timescale of recombination, a conclusion consistent with other experiments^{19–21}. The argument holds also for X^{3-} , which also shows a broad satellite PL (FWHM 1.2 meV).

An obvious feature in Figs 2 and 3 is the marked broadening and increase in intensity in the emission at $V_g \approx 0.1$ V. This looks similar to the broad emission of highly excited dots²². We suggest that the PL is broad at 0.1 V because relaxation of the final state always proceeds quickly, regardless of spin. A new relaxation channel must open at 0.1 V. We know from the capacitance that at 0.1 V the wetting layer begins to fill, and the new channel is therefore an Auger process: electrons in the quantum ring lose energy by promoting an electron in the wetting layer to an unoccupied state at higher energy²³. Auger processes can swap the spins of the two participating electrons. In fact, spin-swapping processes are most likely²⁴. This interpretation is supported by the nature of the PL at higher positive voltages, where the PL to some extent narrows again. The ring electrons can only interact with electrons in the continuum a few meV away from the Fermi energy, so that once the Fermi energy is very high, the Auger rate will decrease. The broadening at 0.1 V is accompanied by a rapid increase in overall intensity. This is partly caused by an increase in capture efficiency, but other factors may be involved. $\square$

Received 20 January; accepted 3 May 2000.

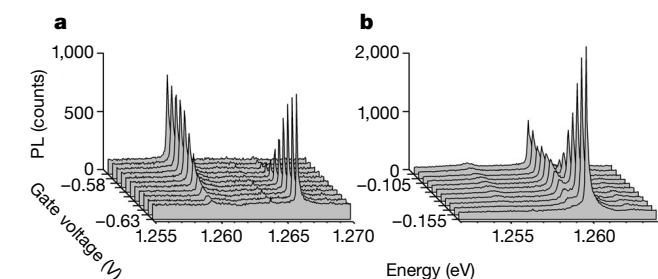


Figure 4 PL of a single quantum ring in the transition from one charge state to the next. **a**, X to X^{1-} ; **b**, X^{1-} to X^{2-} . Panel **b** also shows the emergence of the satellite at 1.255 eV.

1. Tarucha, S., Austing, D. G., Honda, T., van der Hage, R. J. & Kouwenhoven, L. P. Shell filling and spin effects in a few electron quantum dot. *Phys. Rev. Lett.* **77**, 3613–3616 (1996).
2. Miller, B. T. *et al.* Few-electron ground states of charge-tunable self-assembled quantum dots. *Phys. Rev. B* **56**, 6764–6769 (1997).
3. Banin, U., Cao, Y., Katz, D. & Millo, O. Identification of atomic-like electronic states in indium arsenide nanocrystal quantum dots. *Nature* **400**, 542–544 (1999).
4. Bimberg, D., Grundmann, M. & Ledentsov, N. N. *Quantum Dot Heterostructures* (Wiley, Chichester, 1998).
5. Yusa, G. & Sakaki, H. Trapping of photogenerated carriers by InAs quantum dots and persistent photoconductivity in novel GaAs/n-AlGaAs field-effect transistor structures. *Appl. Phys. Lett.* **70**, 345–347 (1997).
6. Finley, J. J. *et al.* Electrical detection of optically induced charge storage in self-assembled InAs quantum dots. *Appl. Phys. Lett.* **73**, 2618–2620 (1998).
7. Lundström, T., Schoenfeld, W., Lee, H. & Petroff, P. M. Exciton storage in semiconductor self-assembled quantum dots. *Science* **286**, 2312–2314 (1999).
8. Bruchez, M. Jr, Moronne, M., Gin, P., Weiss, S. & Alivisatos, A. P. Semiconductor nanocrystals as fluorescent biological labels. *Science* **281**, 2013–2016 (1998).
9. Garcia, J. M. *et al.* Intermixing and shape changes during the formation of InAs self-assembled quantum dots. *Appl. Phys. Lett.* **71**, 2014–2016 (1997).
10. Lorke, A. *et al.* Spectroscopy of nanoscopic semiconductor quantum rings. *Phys. Rev. Lett.* **84**, 2223–2226 (2000).
11. Drexler, H., Leonard, D., Hansen, W., Kotthaus, J. P. & Petroff, P. M. Spectroscopy of quantum levels in charge-tunable InGaAs quantum dots. *Phys. Rev. Lett.* **73**, 2252–2255 (1994).
12. Ashoori, R. C. *et al.* N-Electron ground state energies of a quantum dot in magnetic field. *Phys. Rev. Lett.* **71**, 613–616 (1993).
13. Medeiros-Ribeiro, G., Pikus, E. G., Petroff, P. M. & Efros, A. L. Single-electron charging and Coulomb interaction in InAs self-assembled quantum dot arrays. *Phys. Rev. B* **55**, 1568–1573 (1997).
14. Wojs, A. & Hawrylak, P. Theory of photoluminescence from modulation-doped self-assembled quantum dots in a magnetic field. *Phys. Rev. B* **55**, 13066–13071 (1997).
15. Blackwood, E., Snelling, M. J., Harley, R. T., Andrews, S. R. & Foxon, C. T. B. Exchange interaction of excitons in GaAs heterostructures. *Phys. Rev. B* **50**, 14246–14254 (1994).

16. Bayer, M. *et al.* Electron and hole g factors and exchange interaction from studies of the exciton fine structure in $\text{In}_{0.60}\text{Ga}_{0.40}\text{As}$ quantum dots. *Phys. Rev. Lett.* **82**, 1748–1751 (1999).
17. Miller, B. T. *et al.* Fine structure in the spectrum of the few-electron ground states of self-assembled quantum dots. *Physica B* **249–251** 257–261 (1998).
18. Empedocles, S. A. & Bawendi, M. G. Quantum-confined Stark effect in single CdSe nanocrystallite quantum dots. *Science* **278**, 2114–2117 (1997).
19. Gammon, D., Snow, E. S., Shanabrook, B. V., Katzer, D. S. & Park, D. Homogeneous linewidths in the optical spectrum of a single gallium arsenide quantum dot. *Science* **273**, 87–90 (1996).
20. Sosnowski, T. S. *et al.* Rapid carrier relaxation in $\text{In}_{0.4}\text{Ga}_{0.6}\text{As}/\text{GaAs}$ quantum dots characterized by differential transmission spectroscopy. *Phys. Rev. B* **57**, R9423–R9426 (1998).
21. Toda, Y., Moriwaki, O., Nishioka, M. & Arakawa, Y. Efficient carrier relaxation mechanism in $\text{InGaAs}/\text{GaAs}$ self-assembled quantum dots based on the existence of continuum states. *Phys. Rev. Lett.* **82**, 4114–4117 (1999).
22. Deckel, E. *et al.* Multiexciton spectroscopy of a single self-assembled quantum dot. *Phys. Rev. Lett.* **80**, 4991–4994 (1998).
23. Bockelmann, U. *et al.* Time resolved spectroscopy of single quantum dots: Fermi gas of excitons? *Phys. Rev. Lett.* **76**, 3622–3625 (1996).
24. Ridley, B. K. *Quantum Processes in Semiconductors* (Clarendon, Oxford, 1982).

Acknowledgements

We thank J. P. Kotthaus and N. D. Drew for discussions. This work was supported by the Deutsche Forschungsgemeinschaft and by QUEST, an NSF Science and Technology Center.

Correspondence and requests for materials should be addressed to R.J.W. (e-mail: R.J.Warburton@hw.ac.uk).

Gigantic optical nonlinearity in one-dimensional Mott–Hubbard insulators

H. Kishida*, H. Matsuzaki*, H. Okamoto*†, T. Manabe‡, M. Yamashita§, Y. Taguchi¶ & Y. Tokura¶#

* Department of Advanced Materials Science, Graduate School of Frontier Sciences, University of Tokyo, Tokyo 113-8656, Japan

† Structure and Transformation Group, PRESTO, Tokyo 113-8656, Japan

‡ Graduate School of Human Informatics, Nagoya University, Nagoya 464-8601, Japan

§ Department of Chemistry, Graduate School of Science, Tokyo Metropolitan University, Hachioji, Tokyo 192-0397, Japan

¶ Structure and Transformation Group, PRESTO, Hachioji, Tokyo 192-0397, Japan

¶ Department of Applied Physics, University of Tokyo, Tokyo 113-8656, Japan

Joint Research Center for Atom Technology (JRCAT), Tsukuba, Ibaraki 305-8562, Japan

The realization of all-optical switching, modulating and computing devices is an important goal in modern optical technology. Nonlinear optical materials with large third-order nonlinear susceptibilities ($\chi^{(3)}$) are indispensable for such devices, because the magnitude of this quantity dominates the device performance. A key strategy in the development of new materials with large nonlinear susceptibilities is the exploration of quasi-one-dimensional systems^{1,2}, or ‘quantum wires’—the quantum confinement of electron–hole motion in one-dimensional space can enhance $\chi^{(3)}$. Two types of chemically synthesized quantum wires have been extensively studied: the band insulators of silicon polymers, and Peierls insulators of π -conjugated polymers and platinum halides. In these systems, $\chi^{(3)}$ values of 10^{-12} to 10^{-7} e.s.u. (electrostatic system of units) have been reported^{3–7}. Here we demonstrate an anomalous enhancement of the third-order nonlinear susceptibility in a different category of quantum wires: one-dimensional Mott insulators of $3d$ transition-metal oxides and halides. By analysing the electroreflectance spectra of these com-

pounds, we measure $\chi^{(3)}$ values in the range 10^{-8} to 10^{-5} e.s.u. The anomalous enhancement results from a large dipole moment between the lowest two excited states of these systems.

Since the discovery of high-temperature superconductivity, the dynamics of charge and spin under the influence of strong electron correlation have been extensively investigated in doped and undoped copper oxides⁸. Among them, Sr_2CuO_3 (whose crystal structure is shown in Fig. 1a) is a prototype one-dimensional (1D) Mott insulator with quantum spin liquid⁹. One-dimensional Cu–O chains are composed of CuO_4 quadrilateral structures with shared corner oxygens along the b axis, in which the Cu ion is divalent (spin quantum number $S = 1/2$) and one unpaired electron exists in the $d_{x^2-y^2}$ orbital. In the Cu–O chain, a 1D electronic state is formed by the overlap of the p_x, p_y orbitals of O and the $d_{x^2-y^2}$ orbitals of Cu, as shown in Fig. 1c. Because of large on-site Coulomb repulsion energy U acting on the Cu ions, the Mott–Hubbard gap is opened in the Cu $3d$ band. The occupied O $2p$ band is located between the Cu $3d$ upper Hubbard (UH) band and the Cu $3d$ lower Hubbard (LH) band, as schematically illustrated in Fig. 1e. A strong charge transfer (CT) transition from the O $2p$ valence band to the Cu $3d$ UH band occurs around 1.75 eV, as observed in the spectrum of the imaginary part of dielectric constant (ϵ_2) shown by the broken line in Fig. 2.

Other prototypical 1D Mott insulators investigated here are the halogen (X)-bridged Ni compounds^{10,11}, $[\text{Ni}(\text{chxn})_2\text{X}]Y_2$ (Y, counter ion; chxn, cyclohexanediamine) whose general crystal structure is shown in Fig. 1b. The Ni^{3+} ions and X^- ions are arranged alternately along the b axis, forming a purely 1D electronic state composed of the p_z orbitals of X and the d_{z^2} orbitals of Ni (see Fig. 1d). Four N atoms of amino groups in two chxn molecules coordinate a Ni^{3+} ion in a plane normal to the chain axis b , and

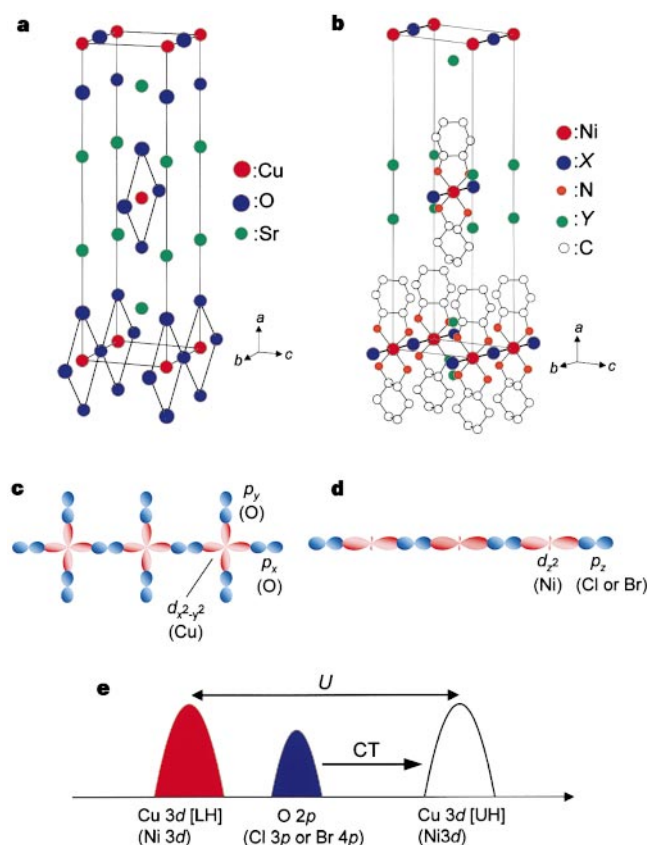


Figure 1 Crystal and electronic structures of the 1D Mott insulators. **a**, Structure of Sr_2CuO_3 . **b**, Structure of $[\text{Ni}(\text{chxn})_2\text{X}]Y_2$. **c**, Configuration of Cu $3d_{x^2-y^2}$ (red) and O $2p_x, 2p_y$ (blue) orbitals in Sr_2CuO_3 . **d**, Configuration of Ni $3d_{z^2}$ (red) and X $2p_z$ (blue) orbitals in $[\text{Ni}(\text{chxn})_2\text{X}]Y_2$. **e**, Schematic electronic structure of the 1D Mott insulators. In **b**, hydrogen atoms are omitted.

produce a strong ligand field. A Ni^{3+} ion is therefore in a low-spin state ($d^7: S = 1/2$) and one unpaired electron exists in the d_{z^2} orbital. This compound is also a Mott insulator (or more exactly, a charge transfer insulator)¹², whose electronic structure is thus analogous to that of the 1D copper oxide (see Fig. 1e). The ϵ_2 spectra of $[\text{Ni}(\text{chxn})_2\text{Br}]\text{Br}_2$ and the isostructural Ni-Cl compounds, $[\text{Ni}(\text{chxn})_2\text{Cl}](\text{NO}_3)_2$ and $[\text{Ni}(\text{chxn})_2\text{Cl}]\text{Cl}_2$, are shown by the broken lines in Fig. 2. Sharp peaks due to the CT transition from the $X p_z$ valence band to the Ni d_{z^2} UH band are observed at 1.3–2.0 eV.

There have been no spectroscopic studies of these 1D Mott insulators to evaluate the nonlinear optical (NLO) response. In order to discuss third-order NLO response in a systematic way, we need to measure $\chi^{(3)}$ spectra over wide energy regions. One of the most useful methods for doing this is electromodulation spectroscopy^{13–15}, which measures electric-field-induced changes of an optical response such as absorption coefficient or reflectivity. This type of nonlinearity is called the d.c. Kerr effect, which is expressed as $\text{Im}\chi^{(3)}(-\omega; 0, 0, \omega)$ (ref. 16). ($\chi^{(3)}$ is defined as $P(\omega_o) = K\epsilon_0\chi^{(3)}(-\omega_o; \omega_a, \omega_b, \omega_c)E(\omega_a)E(\omega_b)E(\omega_c)$ where $P(\omega_o)$ is nonlinear

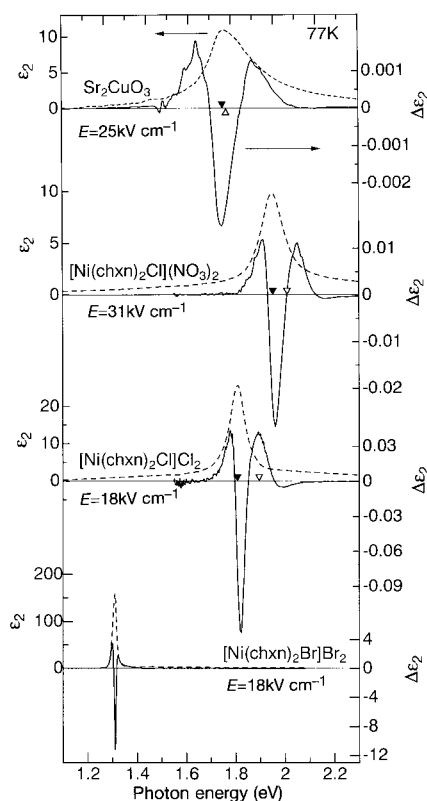


Figure 2 ϵ_2 and $\Delta\epsilon_2$ spectra of Sr_2CuO_3 and $[\text{Ni}(\text{chxn})_2\text{Br}]\text{Br}_2$ and isostructural Ni-Cl compounds. The figure shows the imaginary part of the dielectric constant ϵ_2 (broken line) and its electric-field-induced change $\Delta\epsilon_2$ (solid line) at 77 K of Sr_2CuO_3 , $[\text{Ni}(\text{chxn})_2\text{Cl}](\text{NO}_3)_2$, $[\text{Ni}(\text{chxn})_2\text{Cl}]\text{Cl}_2$ and $[\text{Ni}(\text{chxn})_2\text{Br}]\text{Br}_2$. Single crystals of Sr_2CuO_3 and $[\text{Ni}(\text{chxn})_2\text{X}]\text{Y}_2$ were prepared by the travelling-solvent-floating-zone method²⁰ and the electrochemical method¹⁰, respectively. ϵ_2 spectra were obtained from the polarized reflectivity spectra by using Kramers–Kronig (K–K) transformation. $\Delta\epsilon_2$ spectra were deduced from the electro-reflectance (ER) spectra¹⁴. We adopt the Roessler method²¹ to calculate spectra out of the measured region. In the electro-reflectance measurements, two electrodes spaced ~ 0.5 mm apart are put on the surface of a single-crystal sample, and an alternating voltage (frequency $f = 1$ kHz and amplitude $F \approx 1$ kV) then applied. Both the applied electric field (E) and the polarization of light are set parallel to the 1D chains. The electric-field-induced change (ΔR) of the reflectivity (R) was picked out as the $2f$ component of the reflection light using a lock-in amplifier. By using the K–K transformation, $\Delta R/R$ was converted to $\Delta\epsilon_2$. The energy positions of the first ($\hbar\omega_1$) and second ($\hbar\omega_2$) excited states are indicated by filled triangles and open triangles, respectively.

polarization, $E(\omega_{a,b,c})$ is the electric field, K is the numerical factor and ϵ_0 is the permittivity of vacuum, $\omega_o = \omega_a + \omega_b + \omega_c$.) Here we have applied this technique to the 1D Mott-insulators.

The $\Delta\epsilon_2$ (electric-field-induced change of ϵ_2) spectra polarized along the chain direction are shown for the 1D Mott insulators by the solid lines in Fig. 2. Extremely large $\Delta\epsilon_2$ signals are observed. Among them, the induced change $|\Delta\epsilon_2|$ exceeds 10 for an electric field of $E = 18 \text{ kV cm}^{-1}$ in $[\text{Ni}(\text{chxn})_2\text{Br}]\text{Br}_2$. To discuss the spectral shape in more detail, we show the magnified ϵ_2 and $\Delta\epsilon_2$ spectra of the Ni-Br compound in Fig. 3. $\text{Im}\chi^{(3)}(-\omega; 0, 0, \omega)$ on the right ordinate was calculated from $\Delta\epsilon_2$ using the relation $\Delta\epsilon_2 = 3\text{Im}\chi^{(3)}(-\omega; 0, 0, \omega)E^2$. (In the following, $\chi^{(3)}(-\omega; 0, 0, \omega)$ is expressed simply by $\chi^{(3)}$.) In the $\Delta\epsilon_2$ (or equivalently $\text{Im}\chi^{(3)}$) spectra of the 1D Mott insulators, positive and negative components appear alternately as the photon energy increases.

In the case of inter-band transitions with no electron correlation, it has been theoretically demonstrated that the spectral shape of the electro-absorption $\Delta\epsilon_2$ is almost equal to the third derivative of the linear absorption $d^3\epsilon_2/dE^3$, independent of dimensionality¹⁷. Our $\Delta\epsilon_2$ results, however, cannot be reproduced by $d^3\epsilon_2/dE^3$. But the oscillating structure that we have obtained can be accounted for by assuming two distinct excited states; namely, the one-photon allowed state with odd parity, which is observed as the CT transition in ϵ_2 , and the one-photon forbidden (two-photon allowed) state with even parity. Such an energy level structure is illustrated in Fig. 3 inset. The applied electric field mixes the two excited states. The one-photon allowed state shows a red-shift by the Stark effect, which leads to the increase of absorption at the lower-energy side of the CT transition peak and the decrease at the higher-energy side: that is, the $\Delta\epsilon_2$ spectrum includes the first derivative component of ϵ_2 around the CT transition peak. It is followed by an increase of absorption at the higher-energy side, which comes from the induced absorption of the originally one-photon forbidden state. The superposition of the two components will give an oscillating structure such as that observed in the $\Delta\epsilon_2$ ($\text{Im}\chi^{(3)}$) spectrum.

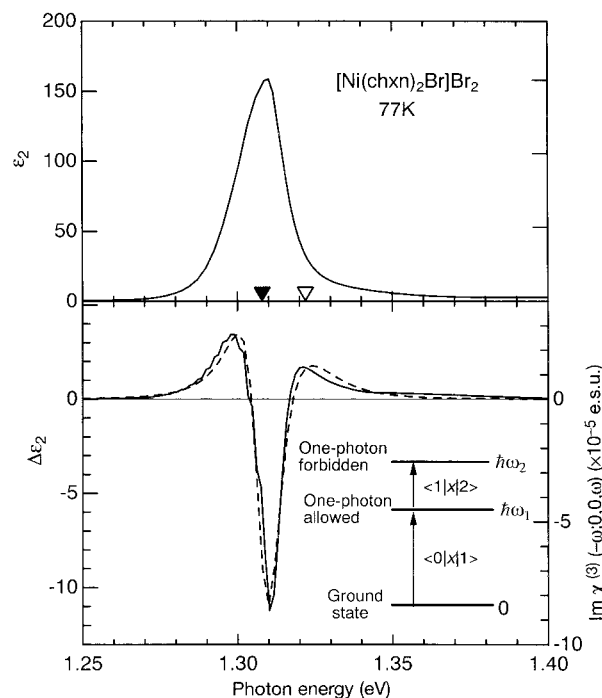


Figure 3 ϵ_2 and $\text{Im}\chi^{(3)}(-\omega; 0, 0, \omega)$ spectra at 77 K of $[\text{Ni}(\text{chxn})_2\text{Br}]\text{Br}_2$. Solid and broken lines are the experimental and calculated results, respectively. In the upper panel, the energy positions of the first ($\hbar\omega_1$) and second ($\hbar\omega_2$) excited states are indicated by a filled triangle and an open triangle, respectively. A schematic energy level structure is inset in the lower panel.

In this case, $\chi^{(3)}$ is given by perturbation theory. The main term dominating the spectral shape of $\chi^{(3)}$ is expressed as follows¹⁶:

$$\chi^{(3)}(-\omega; 0, 0, \omega)_{\text{main}} = \frac{Ne^4}{3\epsilon_0\hbar^3} \frac{\langle 0|x|1\rangle\langle 1|x|2\rangle\langle 2|x|1\rangle\langle 1|x|0\rangle}{(\omega_1 - \omega - i\gamma_1)(\omega_2 - \omega - i\gamma_2)(\omega_1 - \omega - i\gamma_1)} \quad (1)$$

Here $|0\rangle$, $|1\rangle$ and $|2\rangle$ show the ground state, the one-photon allowed excited state and the one-photon forbidden excited state, respectively (see Fig. 3 inset) and $\langle 0,2|x|1\rangle$, the dipole moment between them. $\omega_{1,2}$ and $\gamma_{1,2}$ are the energies and the damping factors for the excited states $|1\rangle$ and $|2\rangle$, respectively. N is the density of Ni or Cu atoms, and ϵ_0 the permittivity of vacuum. To be more exact, $\chi^{(3)}$ is composed of 12 independent terms¹⁶ including equation (1). By using the complete expression of $\chi^{(3)}$, we performed a parameter fitting of the $\text{Im}\chi^{(3)}$ spectrum. The calculated $\text{Im}\chi^{(3)}$ spectrum is given by the broken line in Fig. 3, which reproduces well the experimental result (the solid line). In this fitting procedure, ω_2 , $\gamma_{1,2}$ and $\langle 1|x|2\rangle$ are the fitting parameters, while ω_1 and $\langle 0|x|1\rangle$ are determined from the ϵ_2 spectrum. Similar parameter fittings based on the three-level model have been successfully made on the $\text{Im}\chi^{(3)}$ spectra of Sr_2CuO_3 and the two Ni-Cl compounds. The energy positions of the first ($\hbar\omega_1$) and second ($\hbar\omega_2$) excited states are indicated in Figs 2 and 3 by filled and open triangles, respectively.

As the present results indicate the existence of two distinct levels of excited states, it is natural to consider that both the odd state $|1\rangle$ and the even state $|2\rangle$ are due to CT excitons. But scrutiny of the $\chi^{(3)}$ spectra leads us to suspect the existence of another excited state, in addition to the excited states $|1\rangle$ and $|2\rangle$; the small negative component of $\Delta\epsilon_2$ around 2.0 eV (2.2 eV) in $[\text{Ni}(\text{chxn})_2\text{Cl}]\text{Cl}_2$ ($[\text{Ni}(\text{chxn})_2\text{Cl}](\text{NO}_3)_2$) cannot be explained solely by mixing of the two excited states $|1\rangle$ and $|2\rangle$, but can be explained by assuming another one-photon allowed state located above the state $|2\rangle$. Thus, for the complete understanding of the excited states in the 1D Mott

insulators, it might be necessary to take account of more complicated energy-level structures, although the characteristic spectral shapes of $\chi^{(3)}$ are well explained by the simple three-level model.

In Fig. 4a, the maximum values of $|\text{Im}\chi^{(3)}|$ ($\max|\text{Im}\chi^{(3)}|$) for the 1D Mott insulators are plotted as a function of the optical gap energy ($\hbar\omega_1$). We also show in this figure values of $\max|\text{Im}\chi^{(3)}|$ evaluated by the same experimental method for other 1D semiconductors: a silicon linear polymer (polysilane), which is a band insulator, and π -conjugated polymers and halogen-bridged Pt compounds having Pt-X chains, which are Peierls insulators. As is clearly seen, $\max|\text{Im}\chi^{(3)}|$ values of the 1D Mott insulators ($\sim 10^{-8}$ to 10^{-5} e.s.u.) are larger than those of the other two types of 1D semiconductors ($\sim 10^{-12}$ to 10^{-8} e.s.u.).

We now consider the mechanism of the strong enhancement of $\chi^{(3)}$ in the 1D Mott insulators. In equation (1), based on the three-level model, the important factors dominating $\max|\text{Im}\chi^{(3)}|$ are the transition dipole moments ($\langle 0|x|1\rangle$ and $\langle 1|x|2\rangle$). To extract the contribution of these dipole moments to the material dependence of $\chi^{(3)}$, we plot $\langle 0|x|1\rangle$ and $\langle 1|x|2\rangle$ evaluated by the fitting procedures for the 1D Mott insulators in Fig. 4b, together with the results for other categories of 1D semiconductors—polydihexylsilane (PDHS) and Pt-X compounds ($[\text{Pt}(\text{en})_2][\text{Pt}(\text{en})_2\text{X}_2](\text{ClO}_4)_4$; X = Cl, Br and I). It is significant that the values of $\langle 1|x|2\rangle$ in the 1D Mott insulators are considerably larger than those in PDHS and the Pt-X compounds. Note that $\langle 1|x|2\rangle/\langle 0|x|1\rangle$ exceeds 10. Such an increase of $\langle 1|x|2\rangle$ is an important cause of the enhancement of $\chi^{(3)}$ in the 1D Mott insulators. The value of $\langle 1|x|2\rangle$ gives a crude measure of the spatial extension of the electron-hole wavefunction in the excited states. A large value of $\langle 1|x|2\rangle$ is natural given the large electron transfer energy along the 1D chain, which results from the strong hybridization between the d orbitals of Ni (Cu) ions and the p orbitals of halogen (O) ions. Although the transfer energy is very large, a large optical gap can still exist in the visible or near-infrared region due to the strong electron correlation. These correlations alleviate competing factors (that is, large electron transfer versus large optical gap), and are a great advantage of the 1D Mott insulators as NLO materials.

The other important factors determining the values of $\max|\text{Im}\chi^{(3)}|$ are the damping factors ($\gamma_{1,2}$) and the energy difference between the excited states, $\hbar(\omega_2 - \omega_1)$. The data in Fig. 2 show that the narrower the width of the ϵ_2 peak, the larger the maximum of $\Delta\epsilon_2$. The energy difference $\hbar(\omega_2 - \omega_1)$ evaluated in the 1D Mott insulators (about 0.01–0.14 eV) is relatively small compared to those of PDHS (~ 0.9 eV) and the Pt-X compounds (about 0.2–0.5 eV), which would be also responsible for the enhancement of $\chi^{(3)}$.

When we consider the applications of NLO materials, such as optical switching, modulating and computing devices, a nonlinear susceptibility, $\chi^{(3)}(-\omega; \omega, -\omega, \omega)$, in a transparent region (or equivalently an off-resonant region) is most important. This type of NLO response is observed in other configurations, such as degenerate four-wave mixing (DFWM), pump-and-probe, and two-photon absorption measurements. Here we focus on $\text{Re}\chi^{(3)}(-\omega; \omega, -\omega, \omega)$ at 0.8 eV (1.55 μm), which corresponds to a typical wavelength for optical communication. As for π -conjugated polymers, $|\text{Re}\chi^{(3)}(-\omega; \omega, -\omega, \omega)|$ around the absorption edge (~ 2 eV) was directly measured using DFWM by Bubeck *et al.*⁶ According to the study, $|\text{Re}\chi^{(3)}(-\omega; \omega, -\omega, \omega)|$ for the off-resonant energy of 0.8 eV is much smaller than 10^{-10} e.s.u. Such a measurement of $\chi^{(3)}(-\omega; \omega, -\omega, \omega)$ has not been performed as yet on the 1D Mott insulators presented here. But the values of $|\text{Re}\chi^{(3)}(-\omega; \omega, -\omega, \omega)|$ can be readily calculated by using the parameters ($\omega_{1,2}$, $\gamma_{1,2}$, and $\langle 0,2|x|1\rangle$) obtained from the fitting procedures on the $\chi^{(3)}(-\omega; 0, 0, \omega)$ spectra. The obtained $|\text{Re}\chi^{(3)}(-\omega; \omega, -\omega, \omega)|$ values at 0.8 eV (1.55 μm) are more than 10^{-10} e.s.u. (2.5×10^{-9} e.s.u. for $[\text{Ni}(\text{chxn})_2\text{Br}]\text{Br}_2$, 8.2×10^{-10} e.s.u. for Sr_2CuO_3), though the energy (0.8 eV) is located in the optical gap and is off-resonant to the one-photon allowed state. In particular, $|\text{Re}\chi^{(3)}(-\omega; \omega, -\omega, \omega)|$ in

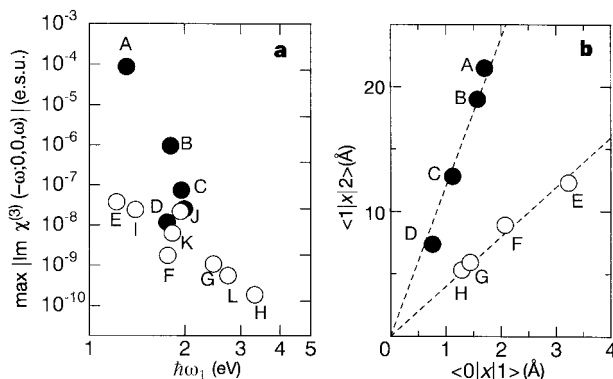


Figure 4 Comparison of the maximum values of $|\text{Im}\chi^{(3)}(-\omega; 0, 0, \omega)|$ and transition dipole moments in various 1D materials. **a**, Maximum values of $|\text{Im}\chi^{(3)}(-\omega; 0, 0, \omega)|$ as a function of the optical gap energies $\hbar\omega_1$. **b**, The relation between the transition dipole moments of $\langle 0|x|1\rangle$ and $\langle 1|x|2\rangle$. Filled circles show the 1D Mott insulators investigated in the present study (A, $[\text{Ni}(\text{chxn})_2\text{Br}]\text{Br}_2$; B, $[\text{Ni}(\text{chxn})_2\text{Cl}]\text{Cl}_2$; C, $[\text{Ni}(\text{chxn})_2\text{Cl}](\text{NO}_3)_2$; D, Sr_2CuO_3), and empty circles the other 1D materials (E, $[\text{Pt}(\text{en})_2][\text{Pt}(\text{en})_2\text{I}_2](\text{ClO}_4)_4$; F, $[\text{Pt}(\text{en})_2][\text{Pt}(\text{en})_2\text{Br}_2](\text{ClO}_4)_4$; G, $[\text{Pt}(\text{en})_2][\text{Pt}(\text{en})_2\text{Cl}_2](\text{ClO}_4)_4$; H, polydihexylsilane (PDHS); I, polyacetylene (PA); J, polydiacetylene (PDA); K, polythienylvinylene (PTV); L, poly(*p*-phenylenevinylene) (PPV). The $|\text{Im}\chi^{(3)}(-\omega; 0, 0, \omega)|$ data of PDHS, PA, PTV and PPV were obtained using unoriented samples. For the other materials, single crystals or oriented films were used. To compare quantitatively the results for the unoriented samples with those for single crystals or oriented films, it is necessary to multiply the $|\text{Im}\chi^{(3)}(-\omega; 0, 0, \omega)|$ values in **a** by 5. For PDHS, $5|\text{Im}\chi^{(3)}(-\omega; 0, 0, \omega)|$ was used for the analysis of the $\chi^{(3)}$ spectrum, taking into consideration the randomness of chain directions. The value of the dipole moment for PDHS plotted in **b** can thus be considered as the value inherent to the polarization parallel to the 1D chain. In **b**, the broken lines are merely guides to the eye.

$[\text{Ni}(\text{chxn})_2\text{Br}]_2$ exceeds 10^{-9} e.s.u., which is, to the best of our knowledge, the largest off-resonant value among the 1D semiconductors. Similar calculations have given $|\text{Re}\chi^{(3)}(-\omega; \omega, -\omega, \omega)|$ values at 0.8 eV of less than 10^{-12} e.s.u. for PDHS and less than 10^{-10} e.s.u. for the Pt-X compounds.

For evaluation of the performance of NLO materials, it is sometimes more practical to compare the figure of merit (FOM), in which the absorption of light is taken into account. The most fundamental form of FOM usually used is $|\chi^{(3)}(-\omega; \omega, -\omega, \omega)|/\alpha(\omega)$ (refs 1, 16, 18, 19). (Here $\alpha(\omega)$ is the linear absorption coefficient.) The above-mentioned $\chi^{(3)}(-\omega; \omega, -\omega, \omega)$ value in the 1D Mott insulators also leads to the following large FOM values ($|\chi^{(3)}|/\alpha$) at 0.8 eV; 2×10^{-12} e.s.u. cm for $[\text{Ni}(\text{chxn})_2\text{Br}]_2$, 8×10^{-13} e.s.u. cm for $[\text{Ni}(\text{chxn})_2\text{Cl}]_2$, 1×10^{-13} e.s.u. cm for $[\text{Ni}(\text{chxn})_2\text{Cl}](\text{NO}_3)_2$, and 7×10^{-13} e.s.u. cm for Sr_2CuO_3 . These values are much larger than that for conjugated polymers, $(5 \pm 3) \times 10^{-14}$ e.s.u. cm, which is known to be almost independent of the photon energy⁶. We therefore conclude that the 1D Mott insulators have third-order optical nonlinearities that are much larger than those of conventional NLO materials, even in an off-resonant condition. The results presented here lead us to expect that the 1D Mott insulators are strong candidates for future NLO materials. □

Received 20 August 1999; accepted 20 April 2000.

1. Auston, D. H. *et al.* Research on nonlinear optical materials: an assessment. *Appl. Opt.* **26**, 211–234 (1987).
2. Sauteret, C. *et al.* Optical nonlinearities in one-dimensional-conjugated polymer crystals. *Phys. Rev. Lett.* **36**, 956–959 (1976).
3. Hasegawa, T. *et al.* Nonlinear optical spectroscopy on one-dimensional excitons in silicon polymer, polysilane. *Phys. Rev. Lett.* **69**, 668–671 (1992).
4. Fann, W.-S. *et al.* Spectrum of $\chi^{(3)}(-3\omega; \omega, \omega, \omega)$ in polyacetylene: an application of the free-electron laser in nonlinear optical spectroscopy. *Phys. Rev. Lett.* **62**, 1492–1495 (1989).
5. Iwasa, Y. *et al.* Nonlinear optical study of quasi-one-dimensional platinum complexes: Two-photon excitonic resonance effect. *Appl. Phys. Lett.* **59**, 2219–2221 (1991).
6. Bubeck, C. *et al.* Resonant degenerate four wave mixing and scaling laws for saturable absorption in thin films of conjugated polymers and Rhodamine 6G. *Chem. Phys.* **154**, 343–348 (1991).
7. Halvorson, C. *et al.* Conjugated polymers with degenerate ground state. The route to high performance third-order nonlinear optical response. *Chem. Phys. Lett.* **200**, 364–368 (1992).
8. Imada, M., Fujimori, A. & Tokura, Y. Metal-insulator transitions. *Rev. Mod. Phys.* **70**, 1039–1263 (1998).
9. Motoyama, N., Eisaki, H. & Uchida, S. Magnetic susceptibility of ideal spin 1/2 Heisenberg antiferromagnetic chain systems, Sr_2CuO_3 and SrCuO_2 . *Phys. Rev. Lett.* **76**, 3212–3215 (1996).
10. Toriumi, K. *et al.* Synthesis and crystal structure of a novel one-dimensional halogen-bridged $\text{Ni}^{\text{III}}\text{-X-Ni}^{\text{III}}$ compounds, $[\text{Ni}(\text{R}, \text{R}'\text{-chxn})_2\text{Br}]_2$. *J. Am. Chem. Soc.* **111**, 2341–2342 (1989).
11. Okamoto, H., Toriumi, K., Mitani, T. & Yamashita, M. Optical and magnetic properties of the halogen-bridged metal complexes modified by hydrogen bondings: $[\text{M}(\text{1R}, 2\text{R-cyclohexanediamine})_2\text{Br}]_2$ ($\text{M} = \text{Pt}, \text{Pd}$, and Ni). *Phys. Rev. B* **42**, 10381–10387 (1990).
12. Okamoto, H. *et al.* Electronic structure of the quasi-one-dimensional halogen-bridged Ni complexes $[\text{Ni}(\text{chxn})_2\text{X}]_2$ ($\text{X} = \text{Cl}, \text{Br}$) and related Ni compounds. *Phys. Rev. B* **54**, 8438–8445 (1996).
13. Seraphin, B. O. Optical field effect in silicon. *Phys. Rev.* **140**, A1716–A1725 (1965).
14. Cardona, M. *Modulation Spectroscopy* (Academic, New York, 1969).
15. Tokura, Y. *et al.* Optical spectra in polydiacetylene crystals substituted with fluorobenzenes. *J. Chem. Phys.* **85**, 99–104 (1986).
16. Butcher, P. N. & Cotter, D. *The Elements of Nonlinear Optics* (Cambridge University Press, 1990).
17. Aspnes, D. E. & Rowe, J. E. Resonant nonlinear optical susceptibility: electroreflectance in the low-field limit. *Phys. Rev. B* **5**, 4022–4030 (1972).
18. Stegeman, G. I. & Wright, E. M. All-optical waveguide switching. *Opt. Quantum Electron.* **22**, 95–122 (1990).
19. Stegeman, G. I. & Torruellas, W. E. Nonlinear materials for information processing and communications. *Phil. Trans. R. Soc. Lond. A* **354**, 745–756 (1996).
20. Suzuura, H. *et al.* Singularities in optical spectra of quantum spin chains. *Phys. Rev. Lett.* **76**, 2579–2582 (1996).
21. Roessler, D. M. Kramers-Kronig analysis of reflection data. *Br. J. Appl. Phys.* **16**, 1119–1123 (1965).
22. Wada, Y. & Yamashita, M. Electric-field effects on the charge-transfer excitation in quasi-one-dimensional halogen-bridged mixed-valence Pt complexes: $[\text{Pt}(\text{ethylenediamine})_2]_2[\text{Pt}(\text{ethylenediamine})_2\text{X}_2](\text{ClO}_4)_4$ ($\text{X} = \text{Cl}, \text{Br}$). *Phys. Rev. B* **42**, 7398–7404 (1990).
23. Phillips, S. D. *et al.* Electroabsorption of polyacetylene. *Phys. Rev. B* **40**, 9751–9759 (1989).
24. Hasegawa, T. *et al.* Analysis of excitonic resonant effect in the third-order nonlinear optical susceptibility of polydiacetylene films. *Synth. Met.* **43**, 3151–3156 (1991).
25. Gelsen, O. M. *et al.* Spectroscopic investigation of the electro-optic nonlinearity in poly(2,5-thienylene vinylene). *J. Appl. Phys.* **71**, 1064–1066 (1992).

Acknowledgements

This work was supported by a grant-in-aid from the Ministry of Education, Science, Sports, and Culture of Japan.

Correspondence and requests for materials should be addressed to H.O. (e-mail: okamoto@ap.t.u-tokyo.ac.jp).

Enantioselective magnetochiral photochemistry

G. L. J. A. Rikken & E. Raupach

Grenoble High Magnetic Field Laboratory, Max Planck Institut für Festkörperforschung/CNRS, BP 166, F-38042 Grenoble, France

Many chemical and physical systems can occur in two forms distinguished solely by being mirror images of each other. This phenomenon, known as chirality, is important in biochemistry, where reactions involving chiral molecules often require the participation of one specific enantiomer (mirror image) of the two possible ones. In fact, terrestrial life utilizes only the L enantiomers of amino acids, a pattern that is known as the 'homochirality of life' and which has stimulated long-standing efforts to understand its origin¹. Reactions can proceed enantioselectively if chiral reactants or catalysts are involved, or if some external chiral influence is present². But because chiral reactants and catalysts themselves require an enantioselective production process, efforts to understand the homochirality of life have focused on external chiral influences. One such external influence is circularly polarized light, which can influence the chirality of photochemical reaction products^{2,13,14}. Because natural optical activity, which occurs exclusively in media lacking mirror symmetry, and magnetic optical activity, which can occur in all media and is induced by longitudinal magnetic fields, both cause polarization rotation of light, the potential for magnetically induced enantioselectivity in chemical reactions has been investigated, but no convincing demonstrations of such an effect have been found^{2–4}. Here we show experimentally that magnetochiral anisotropy—an effect linking chirality and magnetism^{5–7}—can

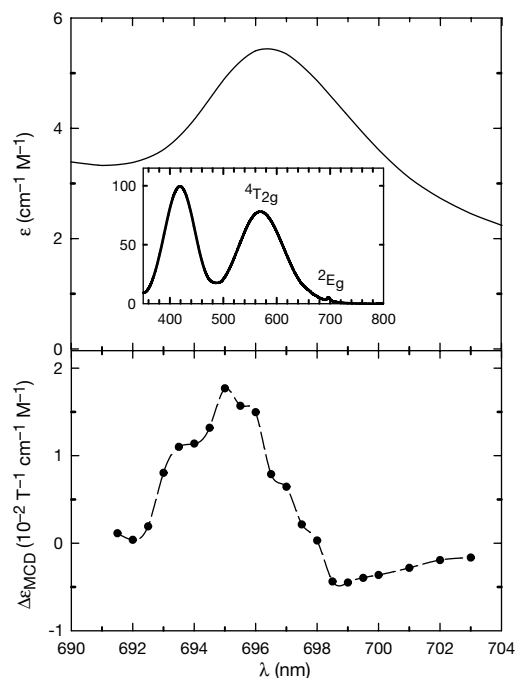


Figure 1 Optical characteristics of the $\text{Cr}(\text{III})$ tris-oxalato complex. Top, absorption spectrum of an aqueous solution of this complex around the ${}^4A_g \rightarrow {}^2E_g$ transition. Inset, the absorption spectrum over the entire visible wavelength range. Bottom, MCD spectrum of an aqueous solution of this complex around the ${}^4A_g \rightarrow {}^2E_g$ transition, obtained by standard polarization modulation techniques.

give rise to an enantiomeric excess in a photochemical reaction driven by unpolarized light in a parallel magnetic field, which suggests that this effect may have played a role in the origin of the homochirality of life.

Interpreting magnetic optical activity as a sign of magnetically induced chirality, Pasteur was the first to search—in vain—for an enantioselective effect of magnetic fields³, followed by many others. As pointed out by Barron, who formulated the symmetry requirements for a reaction to yield a chiral result, enantioselectivity induced by magnetic fields *per se* is not allowed^{8,9}. However, the magnetochiral anisotropy effect^{5–7} can be regarded as a cross-effect between natural optical activity and magnetic optical activity. It can be described by an extra term in the dielectric constant of a

medium^{10–12}, proportional to $\mathbf{k} \cdot \mathbf{B}$, where $\mathbf{k}$ is the wavevector of the light and $\mathbf{B}$ is the magnetic field. The essential features of MCA are (1) the dependence on the relative orientation of $\mathbf{k}$ and $\mathbf{B}$, (2) the dependence on the handedness of the chiral medium (enantioselectivity), and (3) the independence of the polarization state of the light. The parameter $\mathbf{k} \cdot \mathbf{B}$ fulfils the symmetry requirements for a chiral influence⁸.

Enantioselective photochemistry with circularly polarized light (CPL) is well-established^{13,14}. It is based on natural circular dichroism (NCD) that is, a difference in absorption coefficient of the two enantiomers for CPL. By a similar mechanism, one can expect enantioselectivity in photochemical reactions with unpolarized light in a magnetic field owing to magnetochiral dichroism. The largest magnetochiral dichroism is expected in chiral, paramagnetic complexes containing transition-metal or rare-earth ions¹⁵. The Cr(III)tris-oxalato complex has been extensively studied in the context of enantioselective photochemistry with CPL, in a process called photoresolution¹⁶; this complex, which exists in a right- and a left-handed version, is unstable in solution and spontaneously dissociates and re-associates. At equilibrium, it exists as a racemic mixture (equal concentrations of right- and left-handed complex). The dissociation is accelerated by the absorption of light, so under CPL irradiation, the more absorbing enantiomer will dissociate more often, whereas the subsequent re-association yields equal amounts of both enantiomers. This leads to an excess of the less-absorbing enantiomer, the handedness of which depends on the handedness of the CPL. If this photoresolution is much faster than the thermal racemization, the size of the enantiomeric excess (e.e.) in dynamic equilibrium can be shown to be given by $\text{e.e.}_{\text{CPL}} = g_{\text{NCD}}/2$, where $g_{\text{NCD}} \equiv 2(\epsilon_+ - \epsilon_-)/(\epsilon_+ + \epsilon_-)$ is the NCD asymmetry factor¹⁶, and ϵ_+ and ϵ_- are the optical extinction coefficients for left- and right-handed CPL, respectively. As soon as the irradiation stops, the system will return to the racemic state owing to thermal dissociation and random re-association of the complexes.

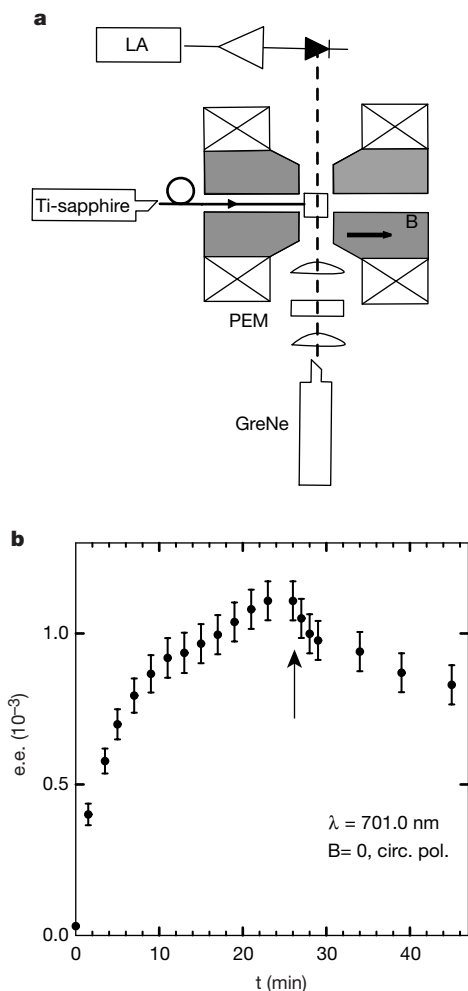


Figure 2 Photoresolution set-up and typical result. **a**, Experimental set-up for determining photoresolution of the Cr(III)tris-oxalato complex. Monochromatic irradiation is performed with a Ti:sapphire laser around 696 nm, with a polarization state that can be selected between linear, circular and unpolarized. The last was obtained after passing the light through 10 m of 1-mm-diameter optical fibre. We have carefully checked that the light leaving the fibre has no net circular component exceeding 10^{-4} . Typical irradiation powers are 100 mW, which are absorbed in 50 μl of 0.2 M aqueous solution of $\text{K}_3\text{Cr}(\text{ox})_3$, with a sample length of 7 mm, kept at 10 °C. Enantiomeric excess (e.e.) detection is done by measuring the NCD at $\lambda = 543.7$ nm (green He–Ne laser, GreNe) using a photo-elastic modulator (PEM) and a lock-in amplifier (LA) and calculating the e.e., using $g_{\text{NCD}}(\lambda = 543 \text{ nm}) = 4.4 \times 10^{-2}$ (ref. 16). Measurements of e.e. are performed within 1 min after the magnetic field and the irradiation were switched off. **b**, Enantiomeric excess as a function of time. Photoresolution was performed by irradiating with 100-mW circularly polarized light at $\lambda = 701$ nm in zero magnetic field. The arrow indicates when the irradiation is stopped, and the system starts to relax to a racemic state. The photoresolution time constant is 6 min, the thermal racemization time constant is 70 min.

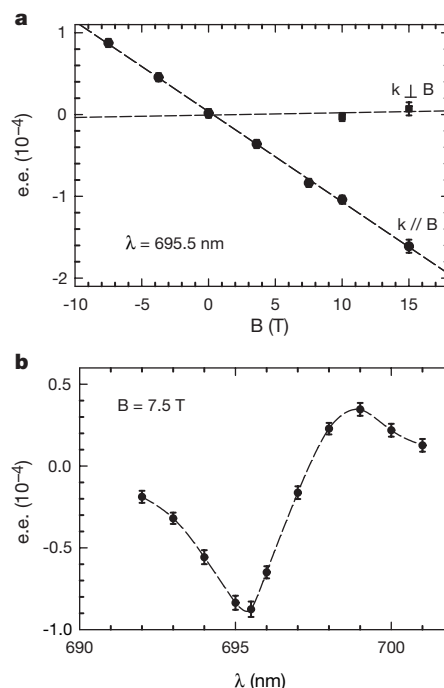


Figure 3 Results of photoresolution with unpolarized light in a magnetic field. **a**, Enantiomeric excess obtained after irradiation with unpolarized light for 25 min at $\lambda = 695.5$ nm, as a function of magnetic field, with an irradiation direction $\mathbf{k}$ either parallel or perpendicular to the magnetic field $\mathbf{B}$. Each point was obtained with a fresh, racemic starting solution. **b**, Enantiomeric excess obtained with the magnetic field of 7.5 T parallel to the irradiation direction, as a function of the irradiation wavelength, with unpolarized light.

Following the same reasoning, we would expect to obtain an enantiomeric excess through magnetochiral anisotropy when one-sidedly irradiating a racemic solution of the Cr(III)tris-oxalato complex with unpolarized light in a magnetic field parallel to the irradiation direction. The handedness of the excess will be determined by the relative orientation of the light and the magnetic field. By a calculation similar to that used for photoresolution with CPL, we find that the magnitude of the excess in dynamic equilibrium is given by $e.e._{MCHA} = g_{MCHA}/2$, where g_{MCHA} is the magnetochiral anisotropy asymmetry factor ($= 2\{\epsilon(\mathbf{k} \uparrow \uparrow \mathbf{B}) - \epsilon(\mathbf{k} \uparrow \downarrow \mathbf{B})\} / \{\epsilon(\mathbf{k} \uparrow \uparrow \mathbf{B}) + \epsilon(\mathbf{k} \uparrow \downarrow \mathbf{B})\}$). We have found that for a given electronic transition, the order of magnitude of g_{MCHA} can be estimated by $(g_{NCD}g_{MCD})/2$ (ref. 7), where $g_{MCD}(B) \equiv 2(\epsilon_+(B) - \epsilon_+(-B)) / (\epsilon_+(B) + \epsilon_+(-B))$ is the magnetic circular dichroism (MCD) asymmetry factor. Both the NCD¹⁷ and the MCD¹⁸ spectra of the Cr(III)tris-oxalato complex are known. In particular, the spin-forbidden transition from the ground state $^4A_{2g}$ to the excited 2E_g state shows fairly large values for both g_{NCD} and g_{MCD} , so we can expect a substantial g_{MCHA} at this transition.

The absorption and MCD spectrum of a racemic solution of the Cr(III)tris-oxalato complex are shown in Fig. 1. We have studied its photoresolution at the $^4A_{2g} \rightarrow ^2E_g$ transition. The set-up is shown in Fig. 2a, together with a typical result with CPL without magnetic field, which serves to show the dynamics of the process (Fig. 2b). Equilibrium is obtained after 20 minutes. After stopping the irradiation, the system relaxes to the racemic state. The photoresolution is found to be much faster than the thermal racemization, and we infer $g_{NCD}(\lambda = 701 \text{ nm}) = 2.2 \times 10^{-3}$.

Figure 3a shows the e.e. obtained in photoresolution with unpolarized light in magnetic field. For fields parallel to the irradiation direction a strictly linear relation, including the sign, is obtained, giving $e.e./B = 1.0 \times 10^{-5} \text{ T}^{-1}$. For perpendicular fields, no significant e.e. is observed. No difference was observed for photoresolution with linearly polarized light or with unpolarized light. These results are in full agreement with the properties of magnetochiral anisotropy, and prove that photochemistry with unpolarized light in magnetic fields can have enantioselectivity. The small magnitude of the observed effect may partly explain why earlier attempts to observe magnetic field effects in enantioselective photochemistry were unsuccessful^{19–21}.

Figure 3b shows the e.e. obtained as a function of the irradiation wavelength, at fixed magnetic field. The dispersive-type lineshape is indicative of so-called A-terms, implying that the influence of the magnetic field on the optical properties is through the Zeeman effect. Such a behaviour was predicted for the magnetochiral anisotropy of the Co(III)tris-oxalato complex, but no detailed prediction for the Cr(III) complex is available¹⁵. As discussed above, the order of magnitude of the e.e. due to magnetochiral anisotropy is estimated to be approximately $(g_{NCD}g_{MCD})/2$. At $\lambda = 701 \text{ nm}$, we have found $g_{MCD}/B = 1.2 \times 10^{-3} \text{ T}^{-1}$ (Fig. 1b) and $g_{NCD} = 2.2 \times 10^{-3}$ (Fig. 2) which yields an estimate of $e.e./B = 1.3 \times 10^{-6} \text{ T}^{-1}$, close to the observed value of $e.e./B = 1.7 \times 10^{-6} \text{ T}^{-1}$.

Magnetochiral anisotropy may also be simulated by a cascading of MCD and NCD⁷. In a magnetic field, MCD leads to a preponderance of one circularly polarized component in the initially unpolarized light. This circularly polarized component will lead to an e.e. through the well-known photoresolution based on g_{NCD} . The true magnetochiral anisotropy will therefore always be accompanied by this so-called cascaded magnetochiral anisotropy. The e.e. that results from such a cascading, averaged over the interaction length of the light with the medium, can—up to moderate absorption and neglecting diffusion and thermal racemization—be shown to be given by

$$e.e._{\text{casc}} \approx \frac{\ln T}{8} g_{MCD} g_{NCD} \quad (1)$$

where T is the optical transmission coefficient. At $\lambda = 701 \text{ nm}$ we calculate that $e.e._{\text{casc}}/B = 5 \times 10^{-7} \text{ T}^{-1}$, which is much smaller than the experimentally observed value, and we therefore conclude that true magnetochiral anisotropy is the dominant process in generating the e.e. observed by us. This is also supported by the fact that $g_{NCD}(\lambda)g_{MCD}(\lambda)$ does not correspond to the observed lineshape of the e.e. In addition, we have determined g_{MCHA} on aqueous solutions of the pure enantiomers (for the experimental setup, see ref. 7), for different concentrations up to that used in the photoresolution measurements. We have found (1) that g_{MCHA} is independent of concentration, which also proves that the true effect dominates, and (2) that within experimental uncertainty $e.e. = g_{MCHA}/2$, which is consistent with the photoresolution mechanism outlined above (E.R. and G.L.J.A.R., manuscript in preparation). The magnitude of the cascaded e.e. could only surpass that of the true e.e. ($\approx (g_{NCD}g_{MCD})/2$) for very small T , that is, strong optical absorption. Under such conditions, equation (1) is no longer valid, as the low intensities in the last part of the sample can no longer ensure that photoresolution is faster than thermal racemization. But, in a two-component system, where MCD and NCD derive their dominant contributions from different electronic transitions, no clear relation between the magnitudes of true and cascaded e.e. can be established. In fact, as both materials with $g_{MCD} > 0.1 \text{ T}^{-1}$ and materials with $g_{NCD} > 0.1$ are known, it is possible that cascaded excesses could be observed in such materials that are much larger than the e.e. obtained in our experiment.

Magnetochiral anisotropy has been suggested as an explanation for the homochirality of life²². So far, the two most likely causes of this homochirality have been considered to be photochemistry with CPL, and the electroweak interaction. The latter, although never experimentally observed in this context, is predicted to lead to extremely small e.e. ($\sim 10^{-17}$), which would require an amplification mechanism that has not yet been identified. Photochemistry with CPL could yield e.e. close to unity. As there is little CPL of natural origin on Earth, this mechanism is assumed to be extraterrestrial. This hypothesis has recently gained some support from the discovery of a L-excess for the amino acids found in a meteorite²³, and from significant amounts of CPL in certain astronomical observations²⁴, albeit not at photochemically active wavelengths. At present it is not known how large a prebiotic e.e. would be required in order to arrive at biological homochirality. We can therefore only wonder if natural magnetic fields—ranging from the Earth's magnetic field of 10^{-4} T , up to the stray field of neutron stars (10^8 T on the surface)²⁵—could, through magnetochiral anisotropy, lead to sufficiently large e.e. Furthermore, issues of spectral, spatial and temporal averaging have to be addressed, similar to the case of photochemistry with CPL. Clearly, the question of the origin of the homochirality of life is far from answered. Our results can only suggest that magnetochiral anisotropy merits consideration in this discussion. They do, however, show convincingly that photochemistry with unpolarized light in a magnetic field is in principle enantioselective. □

Received 4 June 1999; accepted 5 May 2000.

- Cline, D. B. (ed.) *Physical Origin of Homochirality in Life* (American Institute of Physics, New York, 1996).
- Avalos, M. et al. Absolute asymmetric synthesis under physical fields: Facts and fictions. *Chem. Rev.* **98**, 2391–2404 (1998).
- Mason, S. F. in *Circular Dichroism* (eds Nakanishi, K., Berova, N. & Woody, R. W.) Ch. 2 (VCH, New York, 1994).
- Bonner, W. A. Chirality and life. *Orig. Life Evol. Biosphere* **25**, 175–190 (1995).
- Rikken, G. L. J. A. & Raupach, E. Observation of magneto-chiral dichroism. *Nature* **390**, 493–494 (1997).
- Kleindienst, P. & Wagnière, G. Interferometric detection of magnetochiral birefringence. *Chem. Phys. Lett.* **288**, 89–97 (1998).
- Rikken, G. L. J. A. & Raupach, E. Pure and cascaded magnetochiral anisotropy in optical absorption. *Phys. Rev. E* **58**, 5081–5084 (1998).
- Barron, L. D. True and false chirality and absolute asymmetric synthesis. *J. Am. Chem. Soc.* **108**, 5539–5542 (1986).
- Barron, L. D. Can a magnetic field induce absolute asymmetric synthesis? *Science* **266**, 1491–1492 (1994).

10. Portugal, D. L. & Burstein, E. Magneto-spatial dispersion effects on the propagation of electromagnetic radiation in crystals. *J. Phys. Chem. Solids* **32**, 603–608 (1971).
11. Baranova, N. B., Bogdanov, Yu. V. & Zeldovich, B. Ya. Electrical analog of the Faraday effect and other new optical effects in liquids. *Opt. Commun.* **22**, 243–247 (1977).
12. Baranova, N. B. & Zeldovich, B. Ya. Theory of a new linear magnetorefractive effect in liquids. *Mol. Phys.* **38**, 1085–1098 (1979).
13. Rau, H. Asymmetric photochemistry in solution. *Chem. Rev.* **83**, 535–547 (1983).
14. Inoue, Y. Asymmetric photochemical reactions in solution. *Chem. Rev.* **92**, 741–770 (1992).
15. Barron, L. D. & Vrbancich, J. Magneto-chiral birefringence and dichroism. *Mol. Phys.* **51**, 715–730 (1984).
16. Stevenson, K. L. & Verdick, J. F. Partial photoreaction II. Application to some chromium complexes. *Mol. Photochem.* **1**, 271–288 (1969).
17. McCaffery, A. J., Mason, S. F. & Ballard, R. E. Optical rotatory power of coordination compounds. Part III The absolute configurations of trigonal metal complexes. *J. Chem. Soc.* 2883–2892 (1965).
18. McCaffery, A. J., Stephens, P. J. & Schatz, P. N. The magnetic optical activity of d → d transitions. Octahedral chromium(III), cobalt(III), nickel(II) and manganese(II) complexes. *Inorg. Chem.* **6**, 1614–1625 (1967).
19. Pracejus, H. Asymmetrische Synthesen. *Fortschr. Chem. Forsch.* **8**, 493–553 (1967).
20. Bernstein, W. J. The attempted asymmetric synthesis of helices by photolysis under magnetic field. Thesis, Lawrence Berkeley Labs (1972).
21. Teutsch, H. Zur Entstehung optischer Aktivität in der Biosphäre: Ausgewählte Photoreaktionen als asymmetrischen Synthesen. Thesis, Univ. Bremen (1988).
22. Wagnière, G. & Meier, A. Difference in the absorption coefficient of enantiomers for arbitrarily polarized light in a magnetic field—A possible source of chirality in molecular evolution. *Experientia* **39**, 1090–1091 (1983).
23. Engel, M. H. & Macko, S. A. Isotopic evidence for extraterrestrial non-racemic amino acids in the Murchison meteorite. *Nature* **389**, 265–268 (1997).
24. Bailey, J. *et al.* Circular polarization in star-formation regions: Implications for biomolecular homochirality. *Science* **281**, 672–674 (1998).
25. Lyne, A. G. Origins of the magnetic fields of neutron stars. *Nature* **308**, 605–606 (1984).

Acknowledgements

We thank H. Krath for technical assistance; P. Wyder, G. Martinez and B. van Tiggelen for comments on the manuscript; W. A. Bonner for providing important information; L. Barron for discussions; and B. Malezieux for synthesizing the pure enantiomers. The Grenoble High Magnetic Field Laboratory is a “laboratoire conventionné aux universités UJF et INP de Grenoble”.

Correspondence and requests for materials should be addressed to G.L.J.A.R. (e-mail: rikken@polycnrs-gre.fr).

Reduced North Atlantic Deep Water flux to the glacial Southern Ocean inferred from neodymium isotope ratios

Randy L. Rutberg^{†*}, Sidney R. Hemming^{*} & Steven L. Goldstein^{*}

^{*} Lamont-Doherty Earth Observatory and Department of Earth and Environmental Sciences, Columbia University, Palisades, New York 10964, USA

The global circulation of the oceans and the atmosphere transports heat around the Earth. Broecker and Denton¹ suggested that changes in the global ocean circulation might have triggered or enhanced the glacial–interglacial cycles. But proxy data for past circulation taken from sediment cores in the South Atlantic Ocean have yielded conflicting interpretations of ocean circulation in glacial times— $\delta^{13}\text{C}$ variations in benthic foraminifera^{2–6} support the idea of a glacial weakening or shutdown of North Atlantic Deep Water production, whereas other proxies, such as Cd/Ca, Ba/Ca and $^{231}\text{Pa}/^{230}\text{Th}$ ratios, show little change from the Last Glacial Maximum to the Holocene epoch^{7–9}. Here we report neodymium isotope ratios from the dispersed Fe–Mn oxide component of two southeast Atlantic sediment cores. Both cores

show variations that tend towards North Atlantic signatures during the warm marine isotope stages 1 and 3, whereas for the full glacial stages 2 and 4 they are closer to Pacific Ocean signatures. We conclude that the export of North Atlantic Deep Water to the Southern Ocean has resembled present-day conditions during the warm climate intervals, but was reduced during the cold stages. An increase in biological productivity may explain the various proxy data during the times of reduced North Atlantic Deep Water export.

In the present ‘conveyor’ mode, deep water formed in the North Atlantic flows southwards to the circum-Antarctic Southern Ocean. Circumpolar waters flow to the Indian and Pacific oceans, where they upwell and return to the North Atlantic as intermediate and surface waters. The warm and salty northward-flowing surface water releases heat, maintaining mild temperatures in Europe. Whether the conveyor operated during full glacial stages is still an open question, due to the conflicting signals from palaeocirculation proxies in South Atlantic sediments. The importance of ocean circulation in modulating the global climate warrants the use of additional proxies to resolve this conflict.

Neodymium isotope ratios (ϵ_{Nd} ; see Table 1) in sea water and marine precipitates have characteristics that make them powerful palaeoceanographic tracers. They are distinct in different ocean basins, with the highest values in the Pacific ($\epsilon_{\text{Nd}} = 0$ to -4) and the lowest values in the North Atlantic ($\epsilon_{\text{Nd}} \approx -14$), broadly reflecting the ages of the surrounding continents (see refs 10, 11 and references therein). Circum-Antarctic Nd isotope ratios are intermediate, consistent with mixing of Atlantic- and Pacific-derived waters^{11–13}. Like stable-isotope and trace-element proxies, Nd isotope variations in sea water can be directly related to water masses in depth profiles (see, for example, refs 12–16). However, Nd isotope ratios are not measurably fractionated by biological processes or temperature. Therefore, in situations where these effects may cause complications with conventional tracers, the Nd isotope signatures of the source waters are conserved. In the outer layers of Fe–Mn crusts and nodules, Nd isotopes show systematic geographical variations consistent with present-day deep-water circulation¹⁰. However,

Table 1 South Atlantic cores Nd isotope data

Depth in core (cm)	Age (kyr)	$^{143}\text{Nd}/^{144}\text{Nd}$	ϵ_{Nd}
RC11-83			
25	2.5	0.512208 ± 09	-8.39 ± 0.18
126	8.9	0.512180 ± 07	-8.93 ± 0.14
249	14.3	0.512261 ± 12	-7.35 ± 0.23
398	19.7	0.512310 ± 11	-6.40 ± 0.21
600	25.4	0.512296 ± 07	-6.67 ± 0.14
780	33.1	0.512236 ± 21	-7.84 ± 0.41
839	35.8	0.512234 ± 09	-7.88 ± 0.18
999	44.0	0.512185 ± 10	-8.84 ± 0.20
1,101	49.9	0.512204 ± 13	-8.49 ± 0.20
1,251	59.7	0.512260 ± 10	-7.37 ± 0.20
1,285	62.1	0.512271 ± 14	-7.16 ± 0.27
1,381	67.4	0.512224 ± 07	-8.08 ± 0.14
TNO57-6PC			
15	Holocene	0.512222 ± 18	-8.11 ± 0.35
60	Stage 2	0.512432 ± 11	-4.02 ± 0.21

RC11-83 is a dry core from the LDEO Core Repository; TNO57-6PC is a wet core from the ODP Repository at LDEO. Fe–Mn oxide fractions of sediments were leached using hydroxylamine hydrochloride following dissolution of the carbonate fraction with buffered acetic acid. Rare-earth elements (REE) were isolated using TruSpec resin, and Nd was separated using α -hydroxyisobutyric acid and cation resin. The procedural Nd blank is ~ 250 pg, which ranges from 0.2% to 0.8% of the total Nd in each sample, and is thus negligible. Nd was measured by dynamic multicollection as NdO^+ on a VG (Micromass) Sector 54 mass spectrometer at LDEO. $^{143}\text{Nd}/^{144}\text{Nd}$ ratios were normalized to $^{146}\text{Nd}/^{144}\text{Nd} = 0.7219$ and adjusted to a value of 0.511860 for the La Jolla Nd standard. Reported errors are in-run $2\sigma_{\text{mean}}$. The average La Jolla measured values over two measurement periods were 0.511835 ± 22 (2σ external reproducibility, $n = 12$), and 0.511848 ± 23 ($n = 19$), and these (± 0.44 ϵ_{Nd} units) are taken as the errors. Sample RC11-83 at 1,101 cm was measured as Nd^+ by static multicollection on a Micromass P-54 Sector Multicollector ICP-MS at LDEO. The average measured La Jolla value was 0.511862 ± 8 (2σ external reproducibility, $n = 8$). ϵ_{Nd} is the deviation of measured $^{143}\text{Nd}/^{144}\text{Nd}$ ratios from the bulk Earth value of $^{143}\text{Nd}/^{144}\text{Nd} = 0.512638$ in parts per 10^4 . Errors in the $^{143}\text{Nd}/^{144}\text{Nd}$ ratios are in the fifth place after the decimal point.

[†] Present address: Department of Geography, Hunter College of CUNY, 695 Park Avenue, New York, New York 10021, USA.

their slow accumulation rates of a few mm per Myr limit their usefulness to the tracking of ocean circulation changes over glacial–interglacial timescales.

Fe–Mn oxides precipitate from the water column as coatings on biogenic and detrital phases and as particles. To the extent that their Nd isotope ratios record and retain deep-water compositions, they can be directly integrated with high-resolution proxy records in sediment cores. Thus, unlike Fe–Mn nodules and crusts, deep-sea cores allow study of ocean circulation over sub-glacial timescales.

We report downcore Nd isotope data from the dispersed Fe–Mn oxide component of two southeast Atlantic cores. RC11-83 (40° 36' S, 9° 48' E, 4,718 m, ~25 cm kyr⁻¹) is from the Cape basin. TNO57-6PC (42° 55' S, 8° 53' E, 3,750 m, ~15 cm kyr⁻¹) is from south of the Cape basin. Both lie slightly north of the Subpolar Front. RC11-83, extending from the Holocene to the marine isotope stage (MIS) 4–5 transition at ~70 kyr, is one of the most intensively studied cores in the South Atlantic, with a detailed ¹⁴C chronology and stable-isotope record⁴⁶. Variations in benthic foraminiferal $\delta^{13}\text{C}$ (Fig. 1) have been interpreted to indicate a reduced flux of North Atlantic Deep Water (NADW) to the South Atlantic during the last glacial maximum (LGM) and MIS 4.

The Fe–Mn component of samples was isolated using a leaching procedure (Table 1). High-sedimentation-rate cores were sampled in order to minimize possible pore water smearing effects of the Fe–Mn signal over significant time intervals represented in the cores. The integrity of the marine signal was tested in two ways. Firstly, all of the leachates have Quaternary seawater values of $^{87}\text{Sr}/^{86}\text{Sr} = 0.70919\text{--}0.70922$. Coexisting silicate detritus values of 0.717–0.723 (S.L.G., S.R.H., S. Kish and R.L.R., unpublished data) are so much higher than sea water that a tiny amount of detrital contamination of the Fe–Mn leachates would be easily detected. The marine Sr isotope values provide strong evidence that we recovered the seawater signal for Sr. While this does not guarantee a seawater Nd signal, Nd is less soluble than Sr and less likely to experience cross-contamination from detritus. A second test is based on relationships between deep Atlantic seawater Nd isotopes and dissolved silica (Fig. 2a), using silica as an index of mixing between NADW (low SiO₂) and water derived from the Southern Ocean (high SiO₂). They show an excellent correlation, strongly indicating that Nd isotope variations in the deep Atlantic primarily reflect mixing of northern- and southern-derived waters. Nd isotope ratios of Holocene Fe–Mn leachates from the two cores, plotted against the SiO₂ content of the

overlying bottom water¹⁷, lie on this correlation. We conclude from these tests that the Nd isotopes in the Fe–Mn leachates faithfully record the bottom water.

Downcore in RC11-83, Nd isotopes of the Fe–Mn leachates follow the benthic foraminiferal $\delta^{13}\text{C}$ variations (Fig. 1). They vary cyclically with climate stages, with the lowest values in the Holocene and MIS 3 ($\epsilon_{\text{Nd}} \approx -9$), higher values during the LGM and MIS 4 (reaching $\epsilon_{\text{Nd}} \approx -6$), and intermediate values during MIS transitions. Two analyses from the southerly core TNO57-6 show the same Holocene–LGM pattern but a larger difference. Its Holocene ϵ_{Nd} value is similar to that of RC11-83. The deep water at these sites is mainly Circumpolar Deep Water (CDW) from the Antarctic Circumpolar Current (ACC), but the dissolved SiO₂ is lower than Drake Passage CDW (Fig. 2a), indicating the presence of an additional component derived from the North Atlantic. This is fully consistent with the lower ϵ_{Nd} values of the Holocene Fe–Mn leachates compared to Drake Passage CDW (ref. 12). Although TNO57-6 is located only ~2° to the south of RC11-83, its LGM value of $\epsilon_{\text{Nd}} = -4.02$ is significantly higher than that of RC11-83. Both cores lie slightly north of the southeast Atlantic–Circumpolar transition zone, reflected by sharp vertical and latitudinal gradients in dissolved SiO₂ (ref. 17). We speculate that the larger offset of the LGM ϵ_{Nd} values between the cores, in contrast to the similarity in Holocene values, may reflect a northward shift of the transition

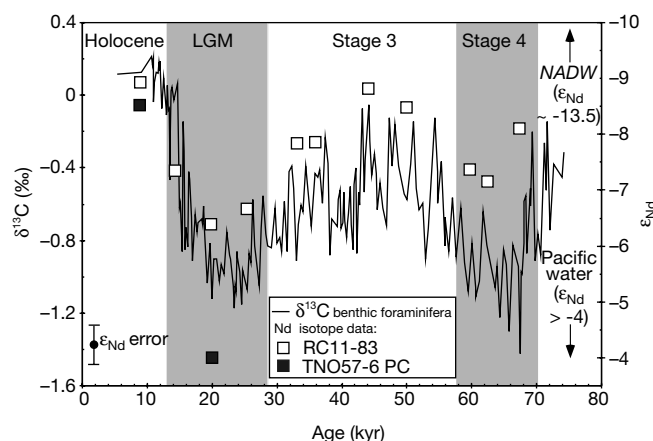


Figure 1 $^{143}\text{Nd}/^{144}\text{Nd}$ ratios of Fe–Mn leachates and benthic foraminiferal $\delta^{13}\text{C}$ versus age. Ages and $\delta^{13}\text{C}$ values for RC11-83 (left axis) are from refs 4 and 6. Holocene and LGM samples for TNO57-6PC are based on the stable-isotope stratigraphy (D. Hodell, personal communication) and not meant to correspond to the precise age. The Nd isotope data (right axis) corroborate the $\delta^{13}\text{C}$ evidence for glacial–interglacial changes in the NADW component at these sites. Core TNO57-6 shows a more extreme excursion towards Pacific-like ϵ_{Nd} values during the LGM.

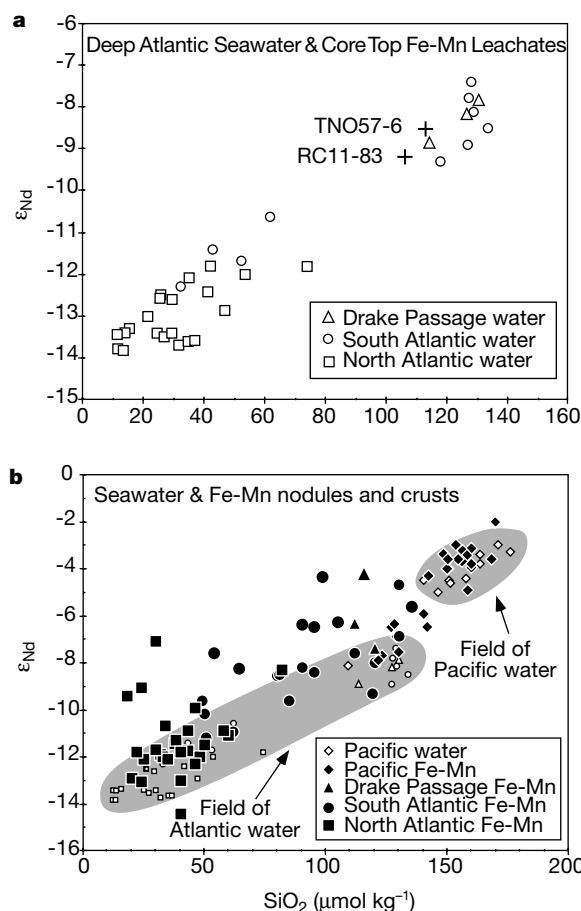


Figure 2 Dissolved silica versus Nd isotope ratios. **a**, ϵ_{Nd} values of dissolved Nd in deep seawater and in South Atlantic core-top Fe–Mn leachates. The leachate data are plotted against bottom-water dissolved silica at the site¹⁷, and fall on the Atlantic seawater trend. **b**, ϵ_{Nd} in deep seawater and in Fe–Mn nodules and crusts. The Fe–Mn samples are plotted against bottom-water dissolved silica at the same sites¹⁷. Each Fe–Mn nodule and crust sample averages over several glacial–interglacial cycles. Most North Atlantic and Pacific Fe–Mn samples fall on the modern seawater trend. South Atlantic samples fall off the trend. Seawater and Fe–Mn data are from the literature.

zone, with southerly core TNO57-6 fully within CDW but RC11-83 still within the transition. The gradient's LGM location may be partly affected by the Agulhas ridge, with ~2,000 m of relief, which restricts deep-water flow into the Cape basin¹⁸.

Atlantic trace-element and stable-isotope records have been interpreted to indicate a glacial shallowing of NADW^{7,19}, or a decrease in the NADW formation rate²⁻⁶ during cold climate stages. The LGM values in both cores are higher than present-day Circumpolar seawater values ($\epsilon_{\text{Nd}} = -7$ to -9). In the southerly core TNO57-6, the LGM value of $\epsilon_{\text{Nd}} = -4.02$ indicates that deep Circumpolar waters had Pacific-like Nd isotope ratios.

The vigour of modern-day ACC flow and the efficiency of vertical mixing are reflected by nearly constant Nd isotope ratios in depth profiles from the Drake Passage¹². Higher Nd isotope ratios in glacial Circumpolar water indicate a larger Pacific Nd component. However, it has been suggested that NADW may have survived as a coherent intermediate-depth water mass in the ACC during full glacial stages²⁰. If the ACC was stratified, then glacial Circumpolar Deep Water might show a reduced NADW signal due to inefficient vertical mixing, rather than a change in the NADW flux to the Southern Ocean. In the modern-day ACC, the low-SiO₂ 'tongue' of NADW can be traced in water column profiles of the Atlantic sector of the Southern Ocean only as far as ~55° E in the western Indian sector¹⁷. Thus, in order for NADW to survive as a coherent water mass around the entire Circumpolar perimeter, an ACC flow regime drastically different from today would be required. A glacial ACC more sluggish than the modern-day is directly at odds with conclusions of a recent study²¹. It is more likely that the reduced influence of Nd derived from the North Atlantic during full glacial stages is a consequence of decreased export of NADW.

Although our data support reduced NADW export to the Southern Ocean during cold climate stages, we have considered alternative possibilities. These include an additional source of Nd to the South Atlantic, a shorter glacial residence time of deep Atlantic Nd (resulting in a smaller NADW-Nd flux reaching the ACC despite a robust NADW-water flux), or changes in the Nd isotope ratio of the NADW or Pacific end-members. We now consider these three alternatives.

Antarctic ice cores show an increased aeolian dust flux derived from Patagonia during MIS 2 and 4²². If the higher particle flux scavenged dissolved Nd from NADW more efficiently during the LGM than today, the strength of the NADW ϵ_{Nd} signal could have been diminished even if the water flux remained constant. The Pacific-like LGM ϵ_{Nd} of southerly core TNO57-6 requires that enhanced scavenging in the Atlantic effectively extracted all of the Nd before the NADW reached the ACC. However, higher glacial dust input to the Pacific would have had the same enhanced scavenging effect on dissolved Nd. If deep water Nd concentrations decreased in both end-members, the Nd isotope ratios of mixtures would be conserved. Alternatively, if the Patagonian dust equilibrated with deep seawater, its high ϵ_{Nd} ($= -3$ to $+8$; ref. 23) could serve as a source of radiogenic Nd in the South Atlantic and ACC. However, in the modern north Pacific, where the aeolian component from China loess ($\epsilon_{\text{Nd}} \approx -10$) dominates the terrigenous input²⁴, Fe-Mn crusts and nodules reflect local bottom water having much higher values ($\epsilon_{\text{Nd}} \approx -4 \pm 2$). In the North Atlantic there is negligible exchange between deep waters and sinking particles 'armoured' by organic material²⁵. Enhanced dust flux has been linked to productivity increases, which would increase the biological packaging of particles, and further inhibit isotopic exchange with sea water²⁶. Thus it is unlikely that enhanced dust fluxes caused the glacial increase of South Atlantic deep-water ϵ_{Nd} values.

Local input from Antarctica could also influence Circumpolar Nd isotope ratios. There are no published Nd isotope analyses of water or sediments near Antarctica, hence Antarctic control on Southern Ocean water cannot be definitively precluded. However, the strong

dissolved SiO₂- ϵ_{Nd} correlation in the modern Atlantic (Fig. 2a) shows that deep-water masses retain characteristic Nd isotope signatures over long travel paths. Nd isotopes in the ACC, intermediate to the North Atlantic and Pacific, are consistent with mixing of these sources, arguing against input from Antarctica as the controlling factor in the Southern Ocean.

Changes in ϵ_{Nd} of the Pacific or Atlantic end-members would also result in changes in the ACC. In the equatorial Pacific, a Fe-Mn crust in which glacial-interglacial timescales are resolvable shows constant ϵ_{Nd} of about -4 through stages 1-6²⁷. Nd isotope ratios in the sampled outer layers of most other Fe-Mn nodules represent averages over 10^5 - 10^6 years, encompassing several glacial cycles. Pacific and most North Atlantic data, plotted against SiO₂ of bottom waters, lie on the modern seawater dissolved SiO₂- ϵ_{Nd} trend (Fig. 2b). In contrast, most South Atlantic data are offset to high ϵ_{Nd} values and thus more Pacific-like than expected from the modern-day bottom-water silica abundance. Thus the available Fe-Mn data indicate that the North Atlantic and Pacific end-members remained stable through the late Pleistocene, and that the South Atlantic, on average, had a larger Pacific component than today.

How sensitive are ACC Nd isotope ratios to NADW flux changes? Mass balance of Nd in the Southern Ocean—assuming an ACC volume of 1.5×10^{17} kg, a Holocene NADW flux of 22 Sv with $\epsilon_{\text{Nd}} = -13.5$ and Nd = 22 pmol l⁻¹, and a Pacific flux of 18 Sv with $\epsilon_{\text{Nd}} = -4.5$ and Nd = 42 pmol l⁻¹ (refs 15, 16, 28, 29)—gives a reasonable ACC water residence time (120 yr) and Nd isotope ratio ($\epsilon_{\text{Nd}} = -8.4$)^{12,13}. Because of the short residence time of Circumpolar water, ACC Nd isotope ratios respond to input changes within hundreds of years. If the NADW flux decreased by 50% during the LGM, then the ϵ_{Nd} of the ACC increases only to -7.0 , insufficient to account for the LGM values (Fig. 1). The present-day NADW value ($\epsilon_{\text{Nd}} = -13.5$) is a mixture of sources from the Labrador Sea ($\epsilon_{\text{Nd}} \approx -20$) and the Denmark Straits and the Norwegian Sea ($\epsilon_{\text{Nd}} \approx -9$)^{14,15}. If the glacial NADW flux was same as today, but the ϵ_{Nd} of NADW increased to -9 (that is, no Labrador Sea component), then the ACC ϵ_{Nd} value increases to -6.4 , again much lower than the LGM value of southerly core TNO57-6. Although the real system is more complicated, some situations are implicit in these scenarios. For example, enhanced loss of NADW Nd in the LGM due to a higher scavenging efficiency follows the same mass balance scenario as a decreased NADW flux. A robust outcome of the mass balance is that an ϵ_{Nd} value of ≈ -5 for the glacial ACC cannot be attained by moderate perturbations to the system. Rather, unless the ϵ_{Nd} of the NADW and Pacific end-members both increased (unsupported by data), or there was an additional high- ϵ_{Nd} source of Nd to the South Atlantic or ACC (unconstrained by data), then the high LGM Nd isotope ratios indicate a drastic decrease in the glacial export of NADW.

The Nd isotope data support interpretations that LGM $\delta^{13}\text{C}$ values of benthic foraminifera in RC11-83 and other South Atlantic cores²⁻⁶ reflect a much-reduced NADW flux to the South Atlantic. However, these conclusions conflict with the no-change interpretations based on Cd/Ca, Ba/Ca and ²³¹Pa/²³⁰Th ratios^{8,9,30}. If the NADW flux decreased, then alternative explanations are required to explain these data. Charles *et al.*⁶ note that the $\delta^{13}\text{C}$ decrease in RC11-83 during the LGM of 1.2‰ (Fig. 1) is more than double that expected from a complete shutdown of NADW. They propose a combined change in thermohaline circulation and an increase in export productivity. Oxidation of the increased amount of organic matter would provide a pool of isotopically light carbon, in which case the low $\delta^{13}\text{C}$ would reflect mixing of organic and bottom-water carbon³¹. The excess CO₂ produced would also enhance CaCO₃ dissolution, which has been shown to cause lowering of Ba/Ca and Cd/Ca ratios³², driving them towards Holocene-like values.

A reduction of the NADW flux and its associated ²³¹Pa load to the glacial Southern Ocean would increase the pool of ²³¹Pa remaining

in the glacial North Atlantic, suggesting an expected glacial pattern of higher North Atlantic and lower South Atlantic $^{231}\text{Pa}/^{230}\text{Th}$ ratios. Because $^{231}\text{Pa}/^{230}\text{Th}$ ratios in sediments co-vary with particle flux as well as with water column ^{231}Pa activity, this pattern might be reversed by changes in particle flux. Enhanced biological activity along glacial North Atlantic continental or ice margins would increase boundary scavenging of ^{231}Pa , leading to a depletion of ^{231}Pa at the middle ocean sampling sites of Yu *et al.*⁹, which might buffer the $^{231}\text{Pa}/^{230}\text{Th}$ ratios to Holocene values (R.F. Anderson, personal communication). In the Southern Ocean, higher glacial diatom productivity near the sample sites (proximal to the Polar Front) would enhance delivery of ^{231}Pa to sediments and counter the effect of reduced import of ^{231}Pa via NADW. Some Southern Ocean studies have interpreted $^{231}\text{Pa}/^{230}\text{Th}$ ratio variations as reflecting productivity changes^{33,34}. A recent study³⁴ concluded that $^{231}\text{Pa}/^{230}\text{Th}$ ratios in Southern Ocean sediments are insensitive to changes in NADW production.

Nd isotope ratios remain unfractionated by productivity changes, but reflect the ages of the continental sources. The simplest explanation for the Nd isotope variations is a decreased NADW flux during cold stages. This, coupled with increased glacial biological productivity, could reconcile the various proxy data for the Southern Ocean. □

Received 18 May 1999; accepted 25 April 2000.

- Broecker, W. S. & Denton, G. H. The role of ocean-atmosphere reorganizations in glacial cycles. *Geochim. Cosmochim. Acta* **53**, 2465–2501 (1989).
- Curry, W. B. & Lohmann, G. P. Reduced advection into Atlantic Ocean deep eastern basins during last glaciation maximum. *Nature* **308**, 317–342 (1983).
- Oppo, D. W. & Fairbanks, R. G. Variability in the deep and intermediate water circulation of the Atlantic Ocean during the past 25,000 years: Northern Hemisphere modulation of the Southern Ocean. *Earth Planet. Sci. Lett.* **86**, 1–15 (1987).
- Charles, C. D. & Fairbanks, R. G. Evidence from Southern Ocean sediments for the effect of North Atlantic deep-water flux on climate. *Nature* **355**, 416–419 (1992).
- Labeysrie, L. *et al.* Hydrographic changes of the Southern Ocean (southeast Indian sector) over the last 230 kyr. *Paleoceanography* **11**, 57–76 (1996).
- Charles, C. D., Lynch-Stieglitz, J., Ninnemann, U. S. & Fairbanks, R. G. Climate connections between the hemisphere revealed by deep sea sediment core/ice core correlations. *Earth Planet. Sci. Lett.* **142**, 19–27 (1996).
- Lea, D. W. & Boyle, E. A. Foraminiferal reconstruction of barium distributions in water masses of the glacial ocean. *Paleoceanography* **5**, 719–742 (1990).
- Boyle, E. A. Cadmium and $\delta^{13}\text{C}$ paleochemical ocean distributions during the Stage 2 glacial maximum. *Annu. Rev. Earth Planet. Sci.* **20**, 245–87 (1992).
- Yu, E.-F., Francois, R. & Bacon, M. P. Similar rates of modern and last glacial ocean thermohaline circulation inferred from radiochemical data. *Nature* **379**, 689–694 (1996).
- Albarède, F. & Goldstein, S. L. World map of Nd isotopes in sea-floor ferromanganese deposits. *Geology* **20**, 761–761 (1992).
- Albarède, F., Goldstein, S. L. & Dautel, D. The neodymium isotopic composition of manganese nodules from the Southern and Indian oceans, the global oceanic neodymium budget, and their bearing on deep ocean circulation. *Geochim. Cosmochim. Acta* **61**, 1277–1291 (1997).
- Piegras, D. J. & Wasserburg, G. J. Isotopic composition of neodymium in waters from the Drake Passage. *Science* **217**, 207–214 (1982).
- Jeandel, C. Concentration and isotopic composition of Nd in the South Atlantic Ocean. *Earth Planet. Sci. Lett.* **117**, 581–591 (1993).
- Stordal, M. C. & Wasserburg, G. J. Neodymium isotopic study of Baffin Bay water: sources of REE from very old terranes. *Earth Planet. Sci. Lett.* **77**, 259–272 (1986).
- Piegras, D. J. & Wasserburg, G. J. Rare earth element transport in the western North Atlantic inferred from Nd isotopic observations. *Geochim. Cosmochim. Acta* **51**, 1257–1271 (1987).
- Piegras, D. J. & Jacobsen, S. B. The isotopic composition of neodymium in the North Pacific. *Geochim. Cosmochim. Acta* **52**, 1373–1381 (1988).
- Levitus, S., Burgett, R. & Boyer, T. *World Ocean Atlas 1994: Nutrients* (US Department of Commerce, Washington DC, 1994).
- Tucholke, B. E. & Embley, R. W. in *Interregional Unconformities and Hydrocarbon Accumulation* (ed. Schlee, J. S.) 145–164 (American Association of Petroleum Geologists, Tulsa, 1984).
- Boyle, E. A. & Keigwin, L. North Atlantic thermohaline circulation during the past 20,000 years linked to high-latitude surface temperature. *Nature* **350**, 35–40 (1987).
- Matsumoto, K. & Lynch-Stieglitz, J. Similar glacial and Holocene deep water circulation inferred from southeast Pacific benthic foraminiferal carbon isotope composition. *Paleoceanography* **14**, 149–163 (1999).
- Pudsey, C. J. & Howe, J. A. Quaternary history of the Antarctic Circumpolar Current: evidence from the Scotia Sea. *Mar. Geol.* **148**, 83–112 (1998).
- Petit, J. R. *et al.* Palaeoclimatological and chronological implications of the Vostok core dust record. *Nature* **343**, 56–58 (1990).
- Basile, I. *et al.* Patagonian origin of glacial dust deposited in East Antarctica (Vostok and Dome C) during glacial stages 2, 4 and 6. *Earth Planet. Sci. Lett.* **146**, 573–589 (1997).
- Jones, C. E., Halliday, A. N., Rea, D. K. & Owen, R. M. Neodymium isotopic variations in North Pacific modern silicate sediment and the insignificance of detrital REE contributions to seawater. *Earth Planet. Sci. Lett.* **127**, 55–66 (1994).

- Jeandel, C., Bishop, J. K. & Zindler, A. Exchange of neodymium and its isotopes between seawater and small and large particles in the Sargasso Sea. *Geochim. Cosmochim. Acta* **59**, 535–547 (1995).
- Fowler, S. W. *et al.* Rapid removal of Chernobyl fallout from Mediterranean surface waters by biological activity. *Nature* **329**, 56–58 (1987).
- Abouchami, W., Goldstein, S. L., Galer, S. J. G., Eisenhauer, A. & Mangini, A. Secular changes of Pb and Nd isotopes in central Pacific seawater as recorded by a Mn crust. *Geochim. Cosmochim. Acta* **61**, 3957–3974 (1997).
- Broecker, W. S. & Peng, T.-H. *Tracers in the Sea* (Eldigio, Palisades, New York, 1982).
- Schmitz, W. *On the World Ocean Circulation: the Pacific and Indian Oceans* (Technical Report WHOI 96–08, Woods Hole Oceanographic Institution, 1996).
- Lea, D. W. A trace metal perspective on the evolution of Antarctic Circumpolar Deep Water Chemistry. *Paleoceanography* **4**, 733–747 (1995).
- Mackensen, A., Hubberten, H. -W., Bickert, T., Fischer, G. & Fütterer, D. K. The $\delta^{13}\text{C}$ in benthic foraminiferal tests of Fontbotia Wuellerstorfi (Schwager) relative to the $\delta^{13}\text{C}$ of dissolved inorganic carbon in Southern Ocean deep water: implications for glacial ocean circulation models. *Paleoceanography* **8**, 587–610 (1993).
- McCorkle, D. C. Evidence of a dissolution effect on benthic foraminiferal shell chemistry: $\delta^{13}\text{C}$, Cd/Ca, Ba/Ca, and Sr/Ca results from the Ontong Java Plateau. *Paleoceanography* **10**, 699–714 (1995).
- Kumar, N. *et al.* Increased biological productivity and export production in the glacial Southern Ocean. *Nature* **378**, 675–679 (1995).
- Asmus, T. *et al.* Variations of biogenic particle flux in the southern Atlantic section of the Subantarctic Zone during the late Quaternary: evidence from $^{231}\text{Pa}_{ex}$ and $^{230}\text{Th}_{ex}$. *Mar. Geol.* **159**, 63–78 (1998).

Acknowledgements

We thank R. Anderson, W. Broecker, L. Burckle and N. Frank for discussions; J. Lynch-Stieglitz for providing the benthic $\delta^{13}\text{C}$ stratigraphy of RC11-83; and D. Hodell for sharing the stratigraphy of TNO57-6PC. R.R. thanks M. Fleisher, G. Hemming, M. Klas-Mendelson, and G. Mandal for help in the laboratory. We also thank A.N. Halliday for comments on the manuscript. This work was partly supported by the NSF and NOAA.

Correspondence and requests for materials should be addressed to S. L. G. (e-mail: steveg@ldeo.columbia.edu).

Mapping the Hawaiian plume conduit with converted seismic waves

X. Li^{*†}, R. Kind^{*†}, K. Priestley[‡], S. V. Sobolev^{*}, F. Tilmann[‡], X. Yuan^{*†} & M. Weber^{*§}

^{*} GeoForschungsZentrum Potsdam, Telegrafenberg, 14473 Potsdam, Germany

[†] Freie Universität, ER Geophysik, Malteserstr. 74-100, 12249 Berlin, Germany

[‡] Bullard Laboratories, Department of Earth Sciences, University of Cambridge, Cambridge CB3 0EZ, UK

[§] Universität Potsdam, Institut für Geowissenschaften, Postfach 601553, 14415 Potsdam, Germany

The volcanic edifice of the Hawaiian islands and seamounts, as well as the surrounding area of shallow sea floor known as the Hawaiian swell, are believed to result from the passage of the oceanic lithosphere over a mantle hotspot^{1–3}. Although geochemical and gravity observations indicate the existence of a mantle thermal plume beneath Hawaii^{4–6}, no direct seismic evidence for such a plume in the upper mantle has yet been found. Here we present an analysis of compressional-to-shear (P-to-S) converted seismic phases, recorded on seismograph stations on the Hawaiian islands, that indicate a zone of very low shear-wave velocity ($< 4 \text{ km s}^{-1}$) starting at 130–140 km depth beneath the central part of the island of Hawaii and extending deeper into the upper mantle. We also find that the upper-mantle transition zone (410–660 km depth) appears to be thinned by up to 40–50 km to the south-southwest of the island of Hawaii. We interpret these observations as localized effects of the Hawaiian plume conduit in the asthenosphere and mantle transition zone with excess temperature of $\sim 300^\circ\text{C}$. Large variations in the transition-zone thickness suggest a lower-mantle origin of the Hawaiian plume similar to the Iceland plume⁷, but our results indicate a 100°C higher temperature for the Hawaiian plume.

Evidence has recently been reported for seismic velocity anoma-

lies immediately above the core–mantle boundary⁸ and in the lower mantle⁹ which might be related to the Hawaiian hotspot. However, no seismic evidence for a thermal plume in the upper mantle beneath Hawaii has been discovered. Lithospheric thickness is estimated (from converted body waves¹⁰) to be ~70 km beneath the island of Oahu, and to be 85–100 km (from Rayleigh wave dispersion) to the northwest of Oahu¹¹ and between the islands of Hawaii and Oahu¹². This lithospheric thickness differs little from the thickness of 80–90-Myr-old undisturbed oceanic lithosphere¹³.

We used seismograms from the IRIS/GEOSCOPE station KIP on the island of Oahu and from a network of six temporary broadband stations on the island of Hawaii (Fig. 1) to look for a possible signature of the Hawaii plume in P-to-S converted waves. The Hawaii stations, hereafter referred to as Hawaii Island Broadband Seismic Network (HIBSN), are operated jointly by the University of Cambridge and the US Geological Survey¹². Receiver functions (RFs) (isolated P-to-S converted waves from seismic velocity contrasts below the seismograph site) from 104 teleseismic recordings from KIP and from 130 recordings from HIBSN are used in this study. The data from all HIBSN stations are summed for each event because these stations are separated by less than 10 km.

The RF data from KIP and HIBSN are presented in Fig. 2. The data-processing steps required to isolate the RFs are well known^{14,15} (see also Fig. 2 legend). The Moho phase is clearly seen in both the

individual and summation traces in Fig. 2a. It is followed at KIP at about 8-s delay time by a strong negative phase. At HIBSN, however, we see a pronounced negative phase (shown in red) at about 14 s, followed by a positive phase. The piercing points (conversion points at a depth) of this negative phase are marked in red in Fig. 1. In Fig. 2b the two conversions from the 410- and 660-km discontinuities are also clearly visible in both individual and summation traces (labelled P410s and P660s). The 660 phase is updoming in a part of Fig. 2b (marked green), with piercing points within the green circle (with a radius of 200 km) shown in Fig. 1. Where both phases have been observed and differential times could be measured in individual traces, these times are indicated in Fig. 1. The thinnest transition zone is observed to the south and southwest of the island of Hawaii, where values 3–5 s smaller than the global average (24 s at 67° epicentral distance) are consistently detected. This observation

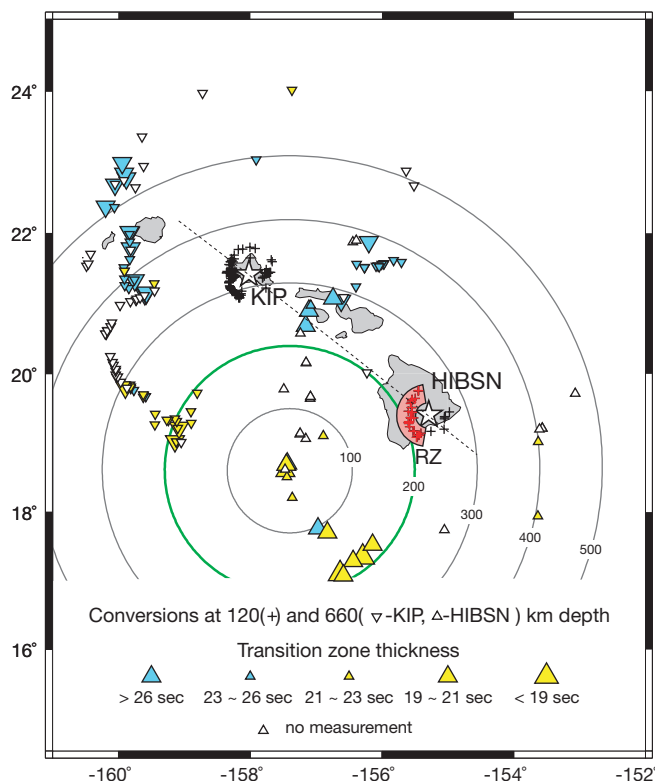


Figure 1 Map of study area. The figure shows the locations of the broadband seismic stations KIP on Oahu and HIBSN on Hawaii, and also the piercing points of P-to-S converted (Ps) seismic phases at 120 and 660 km depth. The colour code and size of the 660-km piercing points indicate the observed differential times between conversions at the 660- and 410-km discontinuities. These numbers are corrected to apply for a teleseismic event at 67° epicentral distance. Yellow piercing points indicate a thinner (hotter), and blue ones a thicker (cooler), transition zone than the average KIP value of 23.2 s (Vinnik *et al.*²⁹ obtained 23.9 s). A concentration of yellow piercing points (hotter material) is found south-southwest of HIBSN with a diameter of about 400 km. Within this region, the green circle marks a confined region of clear updoming of the 660 km converted phase (see Fig. 2b). The red zone in the centre of Hawaii marks a region in the asthenosphere between about 130 and 170 km depth where shear velocity is strongly reduced (see Figs. 2a and 3).

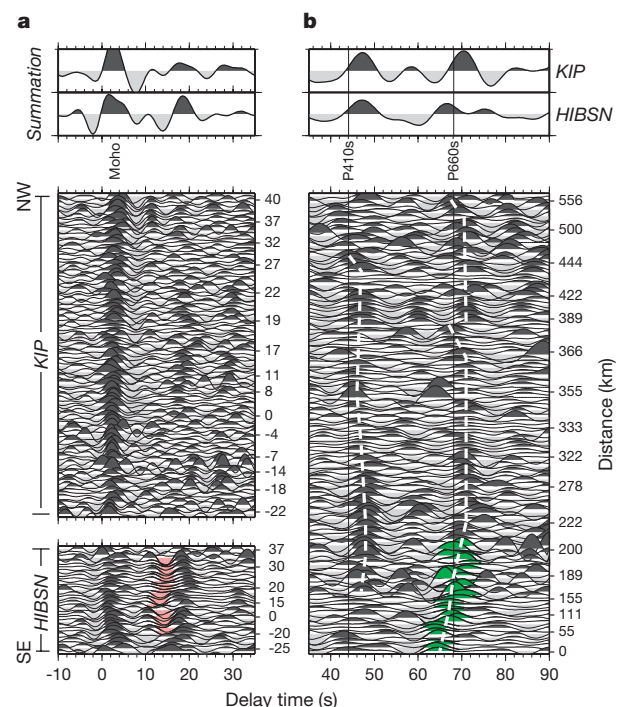


Figure 2 Teleseismic receiver functions (RFs) recorded at KIP and HIBSN. Important processing steps are as follows^{14,15}. (1) Rotation of the vertical, north and east components into the P, SV and SH components of a local ray coordinate system to separate different wave types. Note that we have not rotated into the radial component to avoid projected energy at the P arrival time, which could make the identification of the Moho difficult. (2) Deconvolution of the SV component by the P component in order to make records of different events comparable. (3) Moveout correction using a velocity depth model (the IASP91 model³⁰, reference distance 67°) in order to make records from different epicentral distances comparable (this procedure is identical to the one used in steep angle reflection seismics). Migration as in reflection seismics could also be used¹⁴, but is not applied here. (4) The RFs are filtered with 5-s (a) and 10-s (b) low-pass filters. In a the individual traces are sorted and equally spaced (in order to show all traces) by the projection of their piercing points on a line connecting KIP and HIBSN, which is parallel to the island chain (indicated in Fig. 1). In b the individual traces are sorted and equally spaced by the radial distance of their piercing points at 660 km depth from a hypothetical plume centre at that depth indicated in Fig. 1. Note that the lack of the data below and to the southeast of HIBSN does not allow us to locate this centre precisely. The red phase in a at HIBSN indicates a zone of reduced velocity beneath the island of Hawaii (see Fig. 1 for its location). The 410- and 660-km discontinuities (labelled P410s and P660s) are shown in b. Their arrival times according to the IASP91 global reference model are marked by straight black lines. The dashed white lines show the correlation of the observed conversions. The arrival times of the converted phases have been determined individually at each trace. The part of the 'P660s' marked in green indicates an updoming of the 660-km phase to the south-southwest of Hawaii (within the green circle indicated in Fig. 1). Summation traces of each station are shown at the top of the Figure.

is robust, because it is observed at two different stations with different ray paths in the upper mantle.

Because a 10-km separation change between the 410- and 660-km discontinuities (~ 1 s change in differential time) corresponds to a temperature change of about 70°C (ref. 16), the observed 30–50 km thinning of the transition zone indicates that it has a $200\text{--}350^\circ\text{C}$ higher temperature than the surrounding mantle. The differential times of the summation phases in Fig. 2b are 4 s smaller at HIBSN than at KIP, implying an average excess temperature of about 300°C in the transition zone southwest of HIBSN. This is at least 100°C higher than the temperature estimated at the same depth below Iceland using the same technique⁷. We interpret this temperature anomaly as the thermal signature of the Hawaiian plume conduit in the mantle transition zone. Figure 3 compares the observed uppermost mantle RFs at KIP and HIBSN with the synthetic RFs computed for the Priestley–Tilman surface-wave model¹² and models containing a low-velocity zone (LVZ). The Priestley–Tilman model¹² explains the KIP data reasonably well—although the fit can be further improved (see Supplementary Information)—but not the HIBSN data. The strong negative phase under Hawaii (shown hatched) cannot be explained by crustal multiples (see Supplementary Information), but it is well matched by the LVZ models (Fig. 3). The top of the LVZ is located at 130–140 km depth, well below the expected depth of the lithosphere–asthenosphere boundary beneath the active volcanoes, which is thought to be shallower than that below KIP (less than 100 km depth^{10,12}). The shear velocity in the LVZ is less than 4 km s^{-1} .

An S-velocity of less than 4 km s^{-1} in the ultramafic mantle is inexplicable without the presence of significant partial melt, even if the effect of anelasticity on seismic velocity is considered¹⁷, and the water content of olivine and pyroxene is assumed to be high¹⁸. The observed depth of the LVZ coincides with the estimate of the depth and extent of the melt production region within the plume ($\sim 90\text{--}150\text{ km}$ depth) found by modelling the total melt production, residual depth anomaly, and geoid anomaly⁵; it also coincides with the depth of the deepest region of intensive partial melting

below Hawaii (100–130 km) suggested by the investigation of the primary melt inclusions in olivine crystals¹⁹. The deepest location of partial melting can be expected within the plume conduit where the temperature and water content of the minerals are highest. Therefore, we suggest that the red zone in Fig. 1 indicates the central part of the Hawaiian plume conduit. A few per cent of partial melting is achieved below 130–140 km depth, and this, together with the high temperature and high water content of the plume material, results in the very low observed S-velocity.

At shallower depth within the conduit, the degree of partial melting increases. Melt becomes mobile and leaves the root zone, causing de-watering of the olivine and pyroxene in the root zone. This in turn results in a seismic velocity increase^{18,20}. The olivine de-watering becomes most important at a depth where the mantle temperature exceeds the peridotite dry solidus¹⁸. This depth is close to 60–80 km for the normal asthenospheric temperature, thus providing a possible explanation for the seismic velocity drop observed at this depth beneath the oceans^{18,21}. Due to the higher temperature, the dry solidus temperature is exceeded at a much larger depth within the plume and thus a deeper location of the LVZ can be expected there. The depth to the top of the LVZ in this model depends on the plume temperature, and can therefore be used to estimate the temperature (see Fig. 4). Using the depth dependence of the peridotite dry solidus temperature²² and the depth of the top of the LVZ at 120–140 km, we estimate the temperature within the Hawaiian plume conduit to be higher than $1,600\text{--}1,650^\circ\text{C}$ ($250\text{--}300^\circ\text{C}$ excess temperature). This temperature agrees remarkably well with the petrological estimation of the highest temperature of the Hawaiian primary melts ($1,650^\circ\text{C}$)¹⁹ and with the temperature estimated by geochemical and geophysical modelling^{5,6}. It also agrees with our above estimate of the Hawaiian plume excess temperature in the mantle transition zone.

As according to our interpretation, the 140-km LVZ below the island of Hawaii is related to the hot and undepleted plume conduit material, no similar zone can be expected below Oahu where there is no strong recent volcanism which must be associated with such material. This is in accord with our observations for Oahu (Fig. 3 and Fig. S1 in Supplementary Information).

Thus P-to-S converted waves suggest that the Hawaii plume conduit is located several hundred kilometres to the south-south-west of the island of Hawaii at the depth of the upper-mantle transition zone (see the concentration of yellow piercing points and the green circle in Fig. 1). Unfortunately, due to the absence of ocean-bottom seismographs and the specific distribution of the earthquakes, we cannot constrain the possible plume extent in the continuation of the chain of volcanoes. At asthenospheric depths the plume conduit is located right below the island of Hawaii (see

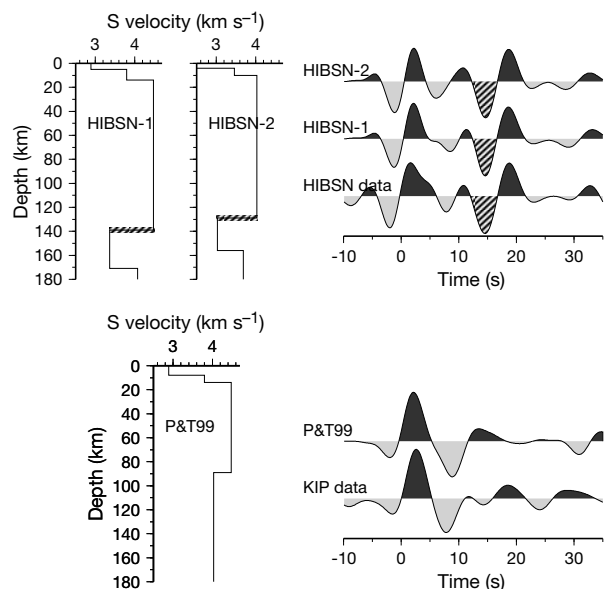


Figure 3 Comparison of observed (summation traces) and synthetic RFs of the uppermost mantle below Oahu (KIP) and Hawaii (HIBSN). Theoretical seismograms of the surface-wave model of Priestley and Tilman¹² (P&T99) explain the observed data at KIP reasonably well, although the fit can be further improved (see Supplementary Information). The hatched phase at HIBSN is well explained by a low-velocity zone (LVZ) between 130 and 170 km depth (see also Supplementary Information for a possible influence of crustal reverberations). The two models HIBSN-1 and HIBSN-2 explain the observations equally well. These models differ in the shear velocity above the LVZ.

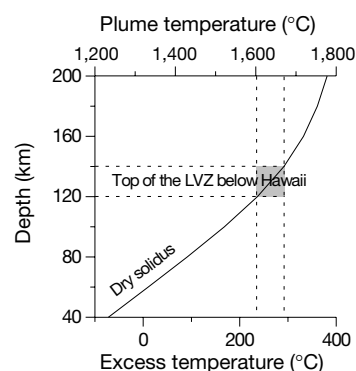


Figure 4 Relation between the depth of the top of the low-velocity zone (LVZ) and plume temperature. The top of the LVZ is associated with the peridotite de-watering effect, which is most pronounced at depths shallower than the depth where the dry peridotite solidus is achieved. The location of the top of the LVZ at 120–140 km below Hawaii corresponds to a plume temperature higher than $1,600\text{--}1,650^\circ\text{C}$ ($250\text{--}300^\circ\text{C}$ excess temperature).

red zone in Fig. 1) in accord with geochemical observations^{23,24}. The conduit is at least 250–300 °C hotter than the surrounding mantle and less than about 400 km in diameter in the transition zone, in line with the absence of the plume signature in a recent global tomographic model²⁵. Due to the large diameter of the Fresnel zone (300 km at 660 km depth) of the converted waves with 10 s period, the actual size of the anomalous zone at 660 km depth may be much smaller than 400 km. The relatively small size of the plume at 660 km depth suggests that the plume originates below this boundary, similar to the finding for the Iceland plume⁷ and consistent with another seismological study indicating lower-mantle origin for another (Bowie) hotspot²⁶. Our estimations suggest, however, that the Hawaii plume temperature is at least 100 °C higher than the temperature of the Iceland plume, which is in accord with petrological observations¹⁹ and recent dynamic models of both plumes^{27,28}.

It is possible that the localized, very-low-velocity zone at asthenospheric depths that we observe is usable as an indicator of a plume conduit. The depth of the top of this zone contains information on the temperature within the plume conduit. □

Received 3 September 1999; accepted 4 May 2000.

1. Wilson, J. T. A possible origin of the Hawaiian Islands. *Can. J. Phys.* **41**, 863–870 (1963).
2. Morgan, W. J. Convection plumes in the lower mantle. *Nature* **230**, 42–43 (1971).
3. Morgan, J. P., Morgan, W. J. & Price, E. Hotspot melting generates both hotspot volcanism and hotspot swell? *J. Geophys. Res.* **100**, 8045–8062 (1995).
4. Hofmann, A. W. Mantle geochemistry: the message from oceanic volcanism. *Nature* **385**, 219–229 (1997).
5. Watson, S. & McKenzie, D. Melt generation by plumes: A study of Hawaiian volcanism. *J. Petrol.* **32**, 501–537 (1991).
6. Ribe, N. M. & Christensen, U. Three-dimensional modelling of plume-lithosphere interaction. *J. Geophys. Res.* **99**, 669–682 (1994).
7. Shen, Y., Solomon, S. C., Bjarnason, I. Th. & Wolfe, C. J. Seismic evidence for a lower-mantle origin of the Iceland plume. *Nature* **395**, 62–65 (1998).
8. Russel, S. A., Lay, T. & Garnero, E. J. Seismic evidence for small-scale dynamics in the lowermost mantle at the root of the Hawaiian hotspot. *Nature* **396**, 255–258 (1998).
9. Ji, Y. & Nataf, H.-C. Detection of mantle plumes in the lower mantle by diffraction tomography: Hawaii. *Earth Planet. Sci. Lett.* **159**, 99–115 (1998).
10. Bock, G. Long-period S to P converted waves and the onset of partial melting beneath Oahu. *Geophys. Res. Lett.* **18**, 869–872 (1991).
11. Woods, M. T., Leveque, J.-J., Okal, E. A. & Cara, M. Two-station measurements of Rayleigh wave group velocity along the Hawaiian Swell. *Geophys. Res. Lett.* **18**, 105–108 (1991).
12. Priestley, K. & Tilmann, F. Shear-wave structure of the lithosphere above the Hawaiian hot spot from two-station Rayleigh wave phase velocity measurements. *Geophys. Res. Lett.* **26**, 1493–1496 (1999).
13. Nishimura, C. & Forsyth, D. Rayleigh wave phase velocities in the Pacific with implications for azimuthal anisotropy and lateral heterogeneities. *Geophys. J. Int.* **94**, 479–501 (1988).
14. Yuan, X., Ni, J., Kind, R., Mechie, J. & Sandvol, E. Lithospheric and upper mantle structure of southern Tibet from a seismological passive source experiment. *J. Geophys. Res.* **102**, 27491–27500 (1997).
15. Kosarev, G. *et al.* Seismic evidence for a detached Indian lithospheric mantle beneath Tibet. *Science* **283**, 1306–1309 (1999).
16. Bina, C. R. & Helffrich, G. Phase transition Clapeyron slopes and transition zone seismic discontinuity topography. *J. Geophys. Res.* **99**, 15853–15860 (1994).
17. Sobolev, S. V. *et al.* Upper mantle temperatures from teleseismic tomography of French Massif Central including effects of composition, mineral reactions, anharmonicity, anelasticity and partial melt. *Earth Planet. Sci. Lett.* **139**, 147–163 (1996).
18. Karato, S. & Jung, H. Water, partial melting and the origin of the seismic low velocity and high attenuation zone in the upper mantle. *Earth Planet. Sci. Lett.* **157**, 193–207 (1998).
19. Sobolev, A. V. & Nikogosian, I. K. Petrology of long-lived mantle plume magmatism: Hawaii, Pacific and Reunion Island, Indian Ocean. *Petrology* **2**, 111–144 (1994).
20. Hirth, G. & Kohlstedt, D. L. Water in the oceanic upper mantle: implications for rheology, melt extraction and the evolution of the lithosphere. *Earth Planet. Sci. Lett.* **144**, 93–108 (1996).
21. Gaherty, J. B., Jordan, T. H. & Gee, L. S. Seismic structure of the upper mantle in a central Pacific corridor. *J. Geophys. Res.* **101**, 22291–22309 (1996).
22. McKenzie, D. & Bickle, M. J. The volume and composition of melt generated by extension of the lithosphere. *J. Petrol.* **29**, 625–679 (1988).
23. Hauri, E. H., Lassiter, J. C. & DePaolo, D. J. Osmium isotope systematics of drilled lavas from Mauna Loa, Hawaii. *J. Geophys. Res.* **101**, 11793–11806 (1996).
24. Hilton, D. R., McMurtry, G. M. & Goff, F. Large variations in vent fluid CO₂/He ratios signal rapid changes in magma chemistry at Loihi seamount, Hawaii. *Nature* **396**, 359–362 (1998).
25. Ekström, G. & Dziewonski, A. M. The unique anisotropy of the Pacific upper mantle. *Nature* **394**, 168–172 (1998).
26. Nataf, H.-C. & VanDecar, J. Seismological detection of a mantle plume? *Nature* **364**, 115–120 (1993).
27. Ribe, N. M. & Christensen, U. R. The dynamical origin of Hawaiian volcanism. *Earth Planet. Sci. Lett.* **171**, 517–531 (1999).
28. Ito, G., Shen, Y., Hirth, G. & Wolfe, C. J. Mantle flow, melting, and dehydration of the Iceland mantle plume. *Earth Planet. Sci. Lett.* **165**, 81–96 (1999).
29. Vinnik, L., Chevrot, S. & Montagner, J. P. Evidence for a stagnant plume in the transition zone? *Geophys. Res. Lett.* **33**, 149–163 (1997).
30. Kennett, B. L. N. *Seismological Tables* (Research School of Earth Sciences, Australian National University, Canberra, 1991).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank IRIS/GEOSCOPE for making the KIP data available to us, P. Dawson and B. Chouet for retrieving the seismic data from the broadband stations on the island of Hawaii, H. Kämpf for discussions, and G. Helffrich for comments on the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft within the International Continental Drilling Project (ICDP) and the GeoForschungsZentrum Potsdam.

Correspondence and requests for materials should be addressed to R.K. (e-mail: kind@gfz-potsdam.de).

A pug-nosed crocodyliform from the Late Cretaceous of Madagascar

Gregory A. Buckley*, Christopher A. Brochu†, David W. Krause‡ & Diego Pol§

* Evelyn T. Stone University College, Roosevelt University, Chicago, Illinois 60605, USA

† Department of Geology, Field Museum of Natural History, Chicago, Illinois 60605, USA

‡ Department of Anatomical Sciences, State University of New York, Stony Brook, New York 11794, USA

§ Division of Paleontology, American Museum of Natural History, New York, New York 10024, USA

Although the image of crocodyliforms as 'unchanged living fossils' is naive, several morphological features of the group are thought to have varied only within narrow limits during the course of evolution¹. These include an elongate snout with an array of conical teeth, a dorsoventrally flattened skull and a posteriorly positioned jaw articulation, which provides a powerful bite force. Here we report an exquisitely preserved specimen of a new taxon from the Late Cretaceous of Madagascar that deviates profoundly from this Bauplan, possessing an extremely blunt snout, a tall, rounded skull, an anteriorly shifted jaw joint and clove-shaped, multicusped teeth reminiscent of those of some ornithischian dinosaurs. This last feature implies that the diet of the new taxon may have been predominantly if not exclusively herbivorous. A close relationship with notosuchid crocodyliforms, particularly *Uruguaysuchus* (Late Cretaceous, Uruguay)² is suggested by several shared derived features; this supports a biogeographical hypothesis that Madagascar and South America were linked during the Late Cretaceous³.

Archosauria Cope 1869

Crocodyliformes Hay 1930

?Notosuchidae Dollo 1924

Simosuchus clarki gen. et sp. nov.

Etymology. Generic name from Greek *simos*, pug-nosed, and Greek *souchos*, the Egyptian crocodile-headed god. Specific name for James M. Clark in recognition of his contributions to crocodyliform systematics.

Holotype. University of Antananarivo UA 8679, complete skull and anterior portion of postcranial skeleton including cervical and anterior dorsal vertebrae, cervical and anterior dorsal osteoderms, and complete pectoral girdle and forelimbs (Fig. 1); discovered by L. L. Randriamaramanana.

Type locality and horizon. Field locality MAD98-17, southeast of the village of Berivotra, Mahajanga Basin, northwestern Madagascar; Maevarano Formation, Upper Cretaceous (Maastrichtian, perhaps late Maastrichtian³).

Diagnosis. Differs from all other crocodyliforms in possessing the following: entire dentition consisting of clove-shaped, multicusped teeth with cusps arranged in single longitudinal row; maxillae do not meet on palatal midline; internarial bar comprising broad ascending premaxillary process; internal nares larger than, and adjacent to, suborbital fenestrae; two ossifications overlying each supratemporal fenestra; quadrate rami project anteroventrally; foramen ovale projects anterolaterally; 14 contiguous quadrilateral osteoderms per mediolateral row in dorsal shield.

Discussion. As indicated by the fusion of neural arches to the vertebral centra, the holotype represents an adult individual⁴, with a skull measuring 12.6 cm anteroposteriorly and 8.2 cm transversely (Fig. 1a–f). The dorsal surface of the skull is lightly sculpted, although the lateral margins of the maxillae are smooth and unsculpted. There are two large palpebral bones over each orbit. The supratemporal fenestrae are distinctly shorter than the orbits,

with an extensive posterior floor formed by the parietal and squamosal. Each is capped by two small, circular ossifications resembling the palpebrals. The squamosal is broad, forming a large shelf overhanging the otic region, with a pronounced posterior projection. The jugal is robust, with an expanded, heavily sculpted ridge below the laterally directed orbit. An antorbital fenestra, found in nearly all notosuchid crocodyliforms⁵, is present and of moderate size. The snout is similar to that of *Comahuesuchus* (Late Cretaceous, Argentina) in being distinctively broad, shortened anteroposteriorly and deep dorsoventrally⁶. Anteriorly, the paired nasals are separated by dorsal extensions of the premaxillae and the paired external nares face forward and laterally.

The quadrate extends anteroventrally, resulting in a jaw articulation that is much further forward than in all other crocodyliforms, with the exception of *Comahuesuchus*. Except for opening and closing, there appears to have been little mobility in jaw movement.

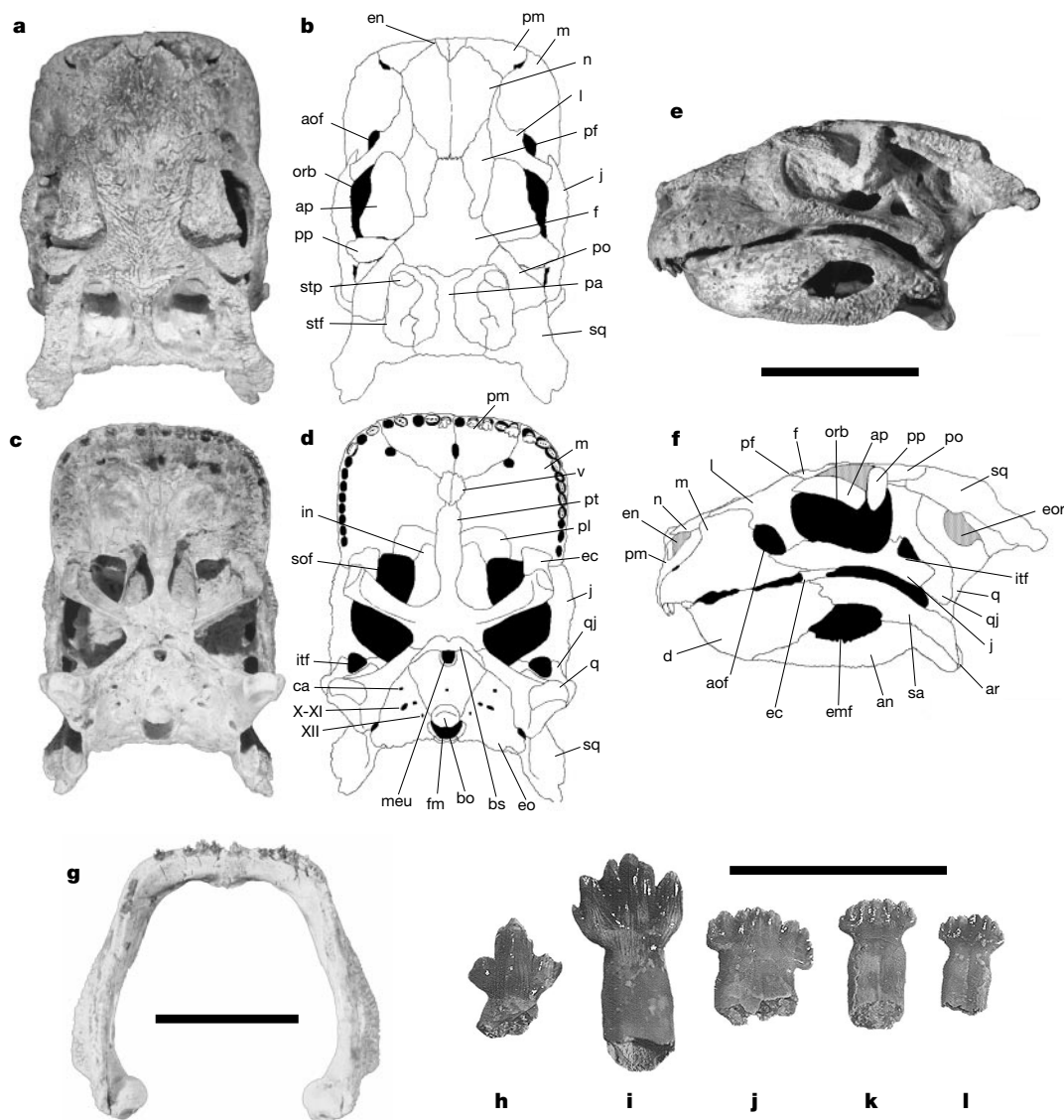


Figure 1 *Simosuchus clarki*, UA 8679 (holotype), from the Upper Cretaceous Maevarano Formation of Madagascar. Skull in dorsal (a, b), ventral (c, d) and lateral (e, f) views; mandible in dorsal view (g); teeth, progressing posteriorly in the tooth row from left to right (h–l). Scale bars, 5 cm (a–g); 1 cm (h–l); an, angular; aof, antorbital fenestra; ap, anterior palpebral; ar, articular; bo, basioccipital; bs, basisphenoid; ca, internal carotid foramen; d, dentary; ec, ectopterygoid; emf, external mandibular fenestra; en, external nares; eo, exoccipital; eor, external otic recess; f, frontal; fm, foramen magnum; in,

internal nares; itf, infratemporal fenestra; j, jugal; l, lacrimal; m, maxilla; meu, medial eustachian foramen; n, nasal; orb, orbit; pa, parietal; pf, prefrontal; pl, palatine; pm, premaxilla; po, postorbital; pp, posterior palpebral; pt, pterygoid; q, quadrate; qj, quadratojugal; sa, surangular; sof, suborbital fenestra; sq, squamosal; stf, supratemporal fenestra; stp, supratemporal ossification; v, vomer; X–XI, opening for posterior cranial nerve; XII, hypoglossal foramen.

The glenoid fossa is narrow anteroposteriorly, suggesting limited propalinal motion. This is in stark contrast to the long jaw articulation in *Notosuchus* (Late Cretaceous, Argentina)⁶ and *Malawisuchus* (Early Cretaceous, Malawi)^{7,8}.

The palatal aspect of the skull is well preserved. The maxillae do not meet along the midline of the palate as in other crocodyliforms, but instead are separated anteriorly by the premaxillae and posteriorly by paired vomers and an extensive anterior projection of the pterygoid. This pterygoid projection also separates the palatines. The shape and position of the vomers in *Simosuchus* are superficially like those of the extant *Melanosuchus*—a diamond-shaped exposure toward the front of the palate⁹. The internal nares open beneath the palatines and the anterior projections of the pterygoid. The occipital condyle is positioned posteroventrally, as in *Notosuchus* and *Malawisuchus*⁸, unlike the posterior orientation typical of nearly all other crocodyliforms.

The mandible is broad and horseshoe-shaped with an extremely short fused symphysis and a laterally expanded flange on the lateroventral margins of the angular (Fig. 1e–g). Although the latter feature is atypical, it is not unique within crocodyliforms, with a similar structure developed in the eusuchian *Mekosuchus* (Cenozoic, Australia and New Caledonia)^{10,11}.

There are 16 tooth positions in each upper jaw quadrant (5 premaxillary and 11 maxillary) and 15 positions in each ramus of the lower jaw. The teeth are unique within crocodyliforms. The dentition of *Simosuchus* is heterodont, although all of the teeth are clove-shaped and multicusped with all of the cusps in a single, longitudinal row. The anteriormost teeth have a large central cusp, with lower accessory cusps (Fig. 1h). Moving back in the tooth row, the cusp number increases and there is greater uniformity in cusp size (Fig. 1i). Still further posteriorly, the crowns are shorter and the teeth become progressively smaller with fewer cusps (Fig. 1j–l). All of the teeth have heavy roots and are strongly compressed buccolingually, with a pronounced constriction at the base of the crown.

Although multicusped teeth are rare in crocodyliforms, they have been reported in *Candidodon* (Early Cretaceous, Brazil)^{12,13}, *Chimaerasuchus* (Early Cretaceous, China)¹⁴, *Malawisuchus* and *Uruguaysuchus*. The teeth of *Uruguaysuchus* are most similar to those of *Simosuchus*, with the former possessing strongly compressed spatulate teeth with a sharp constriction at the base of the crown. Although the cheek teeth of *Uruguaysuchus* have not been described as multicusped², there are distinct cuspules arranged in a single longitudinal row as in *Simosuchus*.

Although unusual, *Simosuchus*, with cranial characters that include an antorbital fenestra smaller than the orbits, a sutured maxillary–premaxillary butt joint, a parietal lacking broad occipital exposure, fenestrated dorsal surface of the quadrate, the presence of two large palpebrals and a flat dorsal surface of the skull table, is indubitably a crocodyliform¹⁵. A phylogenetic analysis, using 22 ingroup taxa and 117 discrete characters, reveals that *Simosuchus* nests firmly within a clade (Fig. 2, node A) consisting of several other small-bodied, short-snouted Gondwanan crocodyliforms, including *Notosuchus* and other putative notosuchids such as *Uruguaysuchus* and *Malawisuchus*. Like all members of this clade, *Simosuchus* possesses a broad and high rostrum, nasals that contact the medial and anterior edges of the lacrimal and a posteroventrally positioned occipital condyle. Within this clade, unambiguous synapomorphies uniting *Simosuchus* with *Uruguaysuchus* and *Malawisuchus* include a long process extending from the postero-lateral edge of the squamosal, a cranioquadrate passage enclosed near the lateral edge of the skull by the quadrate, squamosal and otoccipital, and a retroarticular process that is attenuated and projects posteriorly from the ventral part of the mandible. *Simosuchus* and *Uruguaysuchus* are linked by two unambiguous synapomorphies: internal nares divided by a septum and strongly spatulate posterior teeth. The sister-group relationship of *Simosuchus* and *Uruguaysuchus* corroborates the biogeographic

hypothesis, which is based previously on gondwanatherian mammals¹⁶, abelisaurid theropod dinosaurs¹⁷ and peirosaurid crocodyliforms¹⁸, that Madagascar and South America were physically and biotically linked, perhaps through Antarctica, well into the Late Cretaceous.

Simosuchus and *Uruguaysuchus* both possess posterior teeth with multiple cusps in a single row. This is an ambiguous synapomorphy in this analysis. *Simosuchus*, *Uruguaysuchus* and *Malawisuchus* all share multicusped teeth, but in *Malawisuchus* (as in *Chimaerasuchus*) the cusps are arranged in multiple anteroposterior rows, which was treated as a separate character state. As we treated this character as unordered, neither condition (single-rowed or multiple-rowed multicusped teeth) could be regarded as ancestral for the clade including *Malawisuchus*, *Uruguaysuchus* and *Simosuchus*.

The array of unusual morphological features found in *Simosuchus* leads to several preliminary functional and ecological interpretations. The anterolaterally positioned external nares, together with laterally positioned orbits, suggest that *Simosuchus* was not as well adapted for floating at the surface of an aquatic habitat as are modern crocodylians, in which the external nares and orbits are dorsally positioned¹⁹. Several features suggest that *Simosuchus*, like *Malawisuchus*^{7,8}, may have been an adept head-burrower. These traits, seen in extant head-burrowing vertebrates, include a short, flat, shovel-like snout and deep cranium, a posteroventrally positioned occipital condyle that would orientate the cranium in a more vertical position, a short, underslung lower jaw that would prevent friction from inadvertently opening the jaws during burrowing, and extensive insertion areas for neck musculature (enlarged hypapophyses and elongated neural spines on the cervical vertebrae, expanded squamosal region and occiput)^{20,21}.

The relatively anterior position of the jaw joint and the relative brevity of the mandible suggest that *Simosuchus* emphasized neither force nor speed of the bite typical of modern crocodylians. The geometry of jaw adduction forces may in fact more closely resemble that of turtles, suggesting different modes of food processing²². This, together with marginal, clove-shaped, multicusped teeth adapted for puncturing and shredding, indicates a specialized dietary preference for *Simosuchus* that is not normally displayed by crocodyliforms. Teeth of this form are typically regarded as indicative of

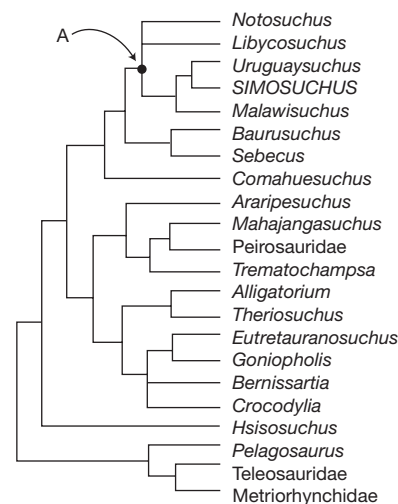


Figure 2 Cladogram showing phylogenetic position of *Simosuchus clarki* based on strict consensus of the six most parsimonious trees (length, 262; Consistency Index (excluding autapomorphies) 0.460; Retention Index, 0.642) generated using PAUP* (version 4.0b2a)²⁷. For character list and taxa/character matrix see Supplementary Information. Trees were rooted with *Protosuchus*, *Orthosuchus* and *Hemiprotosuchus* as outgroups.

an exclusively to predominantly herbivorous diet^{23,24}. In overall morphology the dentition most closely resembles that of some ornithischian dinosaurs, particularly stegosaurs and ankylosaurs, which are also generally regarded as herbivores^{25,26}. Still, there are no modern crocodylian analogues with which to compare the potential function of *Simosuchus*' unique skull and dentition, and the suggestion of an herbivorous diet is speculatively based on dental comparisons with non-crocodylian reptiles. Given the cranial and dental specializations as well as the small size of the adult *Simosuchus*, it can be assumed that *Simosuchus* did not bring down large prey. It is still plausible, however, that the diet may have consisted of arthropods, other invertebrates and potentially small vertebrates such as frogs.

Similarities with ankylosaurid dinosaurs are not limited to dentition. Other features of *Simosuchus* appear to be convergent with those of ankylosaurs, such as the broad compact body, extensive dorsal and ventral shielding, bony protection of the skull above the supratemporal fenestrae and orbits, and a deep cranium with a broad, short snout. A crocodyliform convergent upon an ornithischian dinosaur is intriguing in light of the apparent absence of the latter from the Late Cretaceous of Madagascar. □

Received 8 February; accepted 13 April 2000.

- Langston, W. in *Biology of the Reptilia* Vol. 4 (eds Gans, C. & Parsons, T. S.) 263–284 (Academic, New York, 1973).
- Rusconi, C. Sobre reptiles cretáceos del Uruguay (*Uruguaysuchus* Aznarez, n. g. n. sp.) y sus relaciones con los notosúquidos de Patagonia. *Bol. Inst. Geol. Perf. Uruguay* **19**, 3–64 (1932).
- Krause, D. W. *et al.* The Late Cretaceous vertebrate fauna of Madagascar: implications for Gondwanan paleobiogeography. *GSA Today* **9**, 1–7 (1999).
- Brochu, C. A. Closure of neurocentral sutures during crocodylian ontogeny: implications for maturity assessment in fossil archosaurs. *J. Vertebr. Paleontol.* **16**, 49–62 (1996).
- Pol, D. Basal mesoeucrocodylian relationships: new clues to old conflicts. *J. Vertebr. Paleontol.* **19** (suppl. 3), 69A (1999).
- Bonaparte, J. F. Los vertebrados fósiles de la Formación Rio Colorado, de la ciudad de Neuquén y Cercanías, Cretácico Superior, Argentina. *Rev. Mus. Argent. Cienc. Nat.* **4**, 16–123 (1991).
- Clark, J. M., Jacobs, L. L. & Downs, W. R. Mammal-like dentition in a Mesozoic crocodylian. *Science* **244**, 1064–1066 (1989).
- Gomani, E. M. A crocodyliform from the Early Cretaceous dinosaur beds, northern Malawi. *J. Vertebr. Paleontol.* **17**, 280–294 (1997).
- Medem, F. Osteología craneal, distribución geográfica y ecología de *Melosuchus niger* (Spix) (Crocodylia, Alligatoridae). *Rev. Acad. Colomb. Cienc. Exact. Fis. y Nat.* **12**, 5–19 (1963).
- Balouet, J. C. & Buffetaut, E. *Mekosuchus inexpectatus*, n.g., n.sp., crocodylien nouveau de l'Holocène de Nouvelle Calédonie. *C. R. Acad. Sci.* **304**, 853–856 (1987).
- Willis, P. M. A. New crocodylians from the Late Oligocene White Hunter Site, Riversleigh, north-western Queensland. *Mem. Qld. Mus.* **41**, 423–438 (1997).
- Carvalho, I. S. & Campos, D. A. Um mamífero triconodont do Cretáceo Inferior do Maranhão, Brasil. *An. Acad. Bras. Cienc.* **60**, 437–446 (1988).
- Carvalho, I. S. *Candidodon*: um crocodylo com heterodontia (Notosuchia, Cretáceo Inferior – Brasil). *An. Acad. Bras. Cienc.* **66**, 331–345 (1994).
- Wu, X. -C., Sues, H. -D. & Sun, A. A plant-eating crocodyliform reptile from the Cretaceous of China. *Nature* **376**, 678–680 (1995).
- Benton, M. J. & Clark, J. M. in *The Phylogeny and Classification of the Tetrapods* Vol. 1 (ed. Benton, M. J.) 295–338 (Clarendon, Oxford, 1988).
- Krause, D. W., Prasad, G. V. R., Koenigswald, W. von, Sahni, A. & Grine, F. E. Cosmopolitanism among Gondwanan Late Cretaceous mammals. *Nature* **390**, 504–507 (1997).
- Sampson, S. D. *et al.* Predatory dinosaur remains from Madagascar: implications for the Cretaceous biogeography of Gondwana. *Science* **280**, 1048–1051 (1998).
- Buckley, G. A., Brochu, C. A. & Krause, D. W. Hyperdiversity and the paleobiogeographic origins of the Late Cretaceous crocodyliforms of Madagascar. *J. Vertebr. Paleontol.* **17** (suppl. 3), 35A (1997).
- Jordansky, N. N. in *Biology of the Reptilia* Vol. 4 (eds Gans, C. & Parsons, T. S.) 201–262 (Academic, New York, 1973).
- Gans, C. *Biomechanics: An Approach to Vertebrate Biology* (J. B. Lippincott, Philadelphia, 1974).
- Wake, M. H. in *The Skull* Vol. 3 (eds Hanken, J. & Hall, B. K.) 197–240 (University of Chicago, Chicago, 1993).
- Schumacher, G. -H. in *Biology of the Reptilia* Vol. 4 (eds Gans, C. & Parsons, T. S.) 101–199 (Academic, New York, 1973).
- Hotton, N. III A survey of adaptive relationships of dentition to diet in the North American Iguanidae. *Am. Midl. Nat.* **53**, 88–114 (1955).
- Sues, H. -D. & Reisz, R. R. Origins and early evolution of herbivory in tetrapods. *Trends Ecol. Evol.* **13**, 141–145 (1998).
- Galton, P. M. in *The Dinosauria* (eds Weishampel, D. B., Dodson, P. & Osmólska, H.) 456–483 (Univ. California, Berkeley, 1990).
- Coombs, W. P. & Maryanska, T. in *The Dinosauria* (eds Weishampel, D. B., Dodson, P. & Osmólska, H.) 456–483 (Univ. California, Berkeley, 1990).
- Swofford, D. L. *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4.0b2a. (Sinauer Associates, Sunderland, Massachusetts, 1999).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank B. Rakotosamimanana, A. Rasoamiaramanana, P. Wright, B. Andriamihaja, the staff of the Institute for the Conservation of Tropical Environments, the villagers of Berivotra, and all expedition members for assistance with field work; V. Heisey and J. Holstein for assistance with specimen preparation; B. Demes, C. Forster, N. Rychczynski and S. Zehr for discussion; and the NSF for funding.

Correspondence and requests for materials should be addressed to G.A.B.

Diversity and endemism of the benthic seamount fauna in the southwest Pacific

Bertrand Richer de Forges*†, J. Anthony Koslow‡§ & G. C. B. Poore§

* Centre IRD de Noumea, BP A5, 98848, Noumea Cedex, New Caledonia

‡ CSIRO Marine Research, GPO Box 1538, Hobart, Tasmania 7001, Australia

§ Museum of Victoria, 71 Victoria Crescent, Abbotsford, Victoria 3067, Australia

† These authors contributed equally to this work.

Seamounts comprise a unique deep-sea environment, characterized by substantially enhanced currents and a fauna that is dominated by suspension feeders, such as corals^{1–4}. The potential importance of these steep-sided undersea mountains, which are generally of volcanic origin, to ocean biogeography and diversity was recognized over 40 years ago⁵, but this environment has remained very poorly explored. A review³ of seamount biota and biogeography reported a total of 597 invertebrate species recorded from seamounts worldwide since the *Challenger* expedition of 1872. Most reports, based on a single taxonomic group, were extremely limited: 5 seamounts of the estimated more than 30,000 seamounts in the world's oceans^{4,6} accounted for 72% of the species recorded. Only 15% of the species occurring on seamounts were considered potential seamount endemics. Here we report the discovery of more than 850 macro- and megafaunal species from seamounts in the Tasman Sea and southeast Coral Sea, of which 29–34% are new to science and potential seamount endemics. Low species overlap between seamounts in different portions of the region indicates that the seamounts in clusters or along ridge systems function as 'island groups' or 'chains,' leading to highly localized species distributions and apparent speciation between groups or ridge systems that is exceptional for the deep sea. These results have substantial implications for the conservation of this fauna, which is threatened by fishing activity⁷.

Whereas previous studies of the seamount fauna have often focused on particular taxa, the present study sought to describe the benthic community as a whole and enlisted broad taxonomic support (see Acknowledgements). In all, 516 species of fish and macro-invertebrates were obtained from 6 seamounts along the Norfolk Ridge, 108 from 4 seamounts on the Lord Howe Rise, and 297 from 14 seamounts south of Tasmania (Table 1, Fig. 1). Thirty-six per cent of species from the Norfolk Ridge seamounts were new to science and not known from sampling of the open seafloor and are therefore potential endemic species, along with 31% of species from the Lord Howe seamounts and between 16 and 33% from the Tasmanian seamounts.

Over the range of sampling carried out on Norfolk and Lord Howe Ridges, we found a linear relationship between number of species recorded from each seamount and the number of samples obtained there (Fig. 2). This relationship implies that the greater number of species obtained from seamounts on Norfolk Ridge is

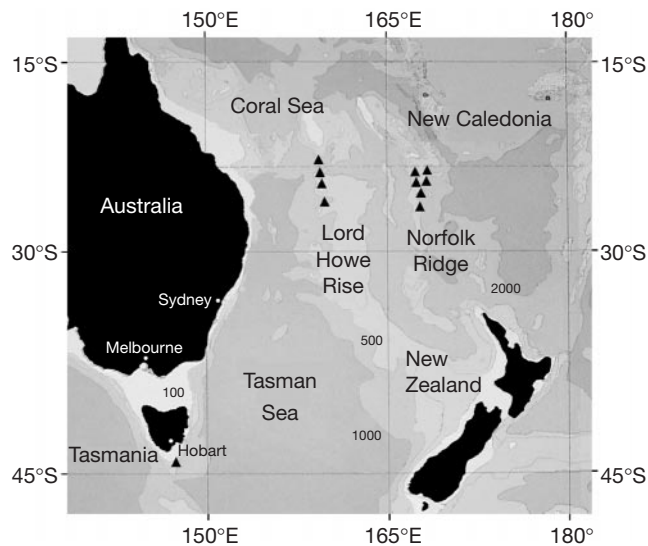


Figure 1 Location of seamount sampling sites (triangles) and topographic and oceanographic features mentioned in the text. Depths are given in fathoms (1 fathom = 1.83 m).

owing to greater sampling effort there rather than greater species richness. The lack of an asymptote in the relationship implies that the number of species present on these seamounts is far greater than the number currently recorded. Species richness on the Tasmanian seamounts appears to be significantly greater, but it is unclear whether this is due to greater diversity or is an artifact of sampling over a greater number of seamounts, albeit within a limited area. Although sampling on the Tasmanian seamounts extended deeper than on Lord Howe Ridge and the Norfolk Rise, sampling in the northern Tasman Sea extended into shallower depths (Table 1). Most species on the Tasmanian seamounts exhibited a broad depth distribution and few were found only at the greatest depths sampled⁷. High species diversity has been previously noted off southeastern Australia^{8,9}.

The coefficient of community (C) between seamounts decreased significantly from a median value of 0.21 (range 0.08–0.49) for adjacent seamounts on the same ridge to a median of 0.04 (range 0–0.17) for seamounts on different ridge systems within the northern sector of the Tasman Sea and southeast Coral Sea (that is, the Norfolk Ridge and Lord Howe Rise) (analysis of variance (ANOVA), $F_{1,43}$ (C) = 46.56, $P < 0.001$) (Fig. 3). Just four species from the Tasmanian seamounts are known from the seafloor around New Caledonia (the bryozoan *Haswelliporina multiaviculata*, the bivalve *Acesta* sp., the barnacle *Hexelasma* sp., and the hydroid *Symplectoscyphus commensalis*), but none of these was found on

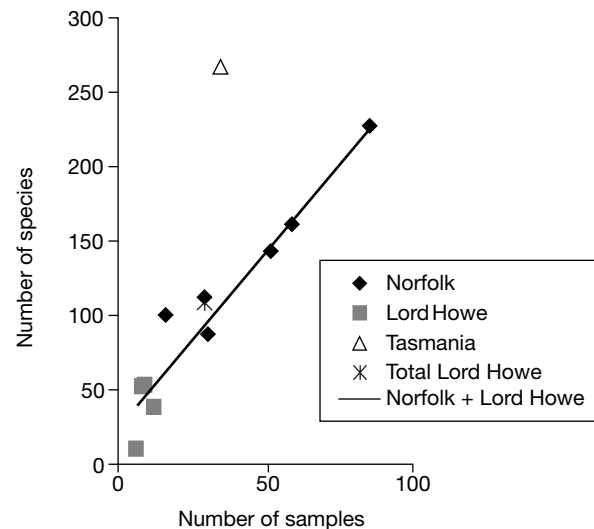


Figure 2 Relationship between the number of species recorded at each seamount site and the number of samples obtained from that site.

seamounts in that region ($C = 0$). Thus there are significant differences in species composition of the seamount communities between ridge systems at a similar latitude and separated by ~1,000 km across the northern Tasman Sea and southeast Coral Sea, and a complete changeover of species between the northern and southern sectors of the Tasman Sea. C did not decline significantly with distance on smaller scales (tens to several hundred kilometres) (Fig. 3).

Although the region around southern Australia and New Caledonia is noted for its relatively isolated faunas^{10,11}, the complete lack of affinity between the seamount communities across this area appears remarkable. For comparison, about 60% of the 93 species of non-mesopelagic decapod crustaceans (Anomura, Caridea, Palinura, Brachyura and Astacidea) recorded from depths greater than 200 m on the continental slope off southeastern Australia (south of ~35° S latitude) are recorded from the Indo-West Pacific region as well (G.C.B.P., unpublished data). This mirrors the general pattern for decapods on the continental shelf in this region, although isopod affinities are mostly with Antarctic and deep-sea faunas⁹. Thus, the isolation of the seamount faunas between the southern and northern Tasman Sea contrasts markedly with the generally strong affinity displayed by the soft-sediment slope fauna between these two regions.

It was also striking to find such little overlap in community composition in the deep sea between sites sharing the same habitat

Table 1 Location and depth range of seamounts sampled

Seamount	Latitude	Longitude	Seamount depth range (m)	Sampling depth range (m)	Number of samples
Norfolk Ridge					
Aztec	24° 07'	168° 45'	55–1,000	57–800	16
Sponge	24° 55'	168° 20'	450–2,000	494–920	84
Jumeau East	23° 42'	168° 17'	250–1,500	380–950	30
Jumeau West	23° 40'	168° 00'	200–1,000	230–730	51
Kaimon Maru	23° 26'	168° 05'	400–2,000	223–600	58
Stylaster	23° 40'	167° 40'	400–1,000	418–970	29
Lord Howe					
Argo	23° 13'	159° 35'	150–2,000	178–300	6
Kelso	23° 51'	159° 24'	150–2,000	160–840	12
Nova	22° 28'	159° 15'	250–2,000	277–930	9
Capel	24° 50'	159° 35'	200–2,000	223–540	8
Tasmanian*	44° 12' – 44° 24'	146° 59' – 147° 23'	660–2,000	660–1,700	34

* The Tasmanian seamounts are aggregated owing to limited sampling on 14 individual seamounts.

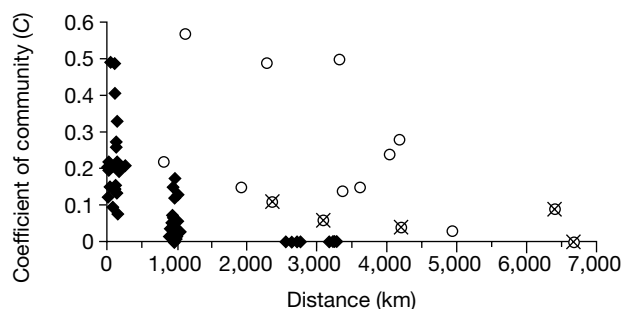


Figure 3 The coefficient of community (C) in relation to distance. C is shown between seamount sites (inverted solid diamonds), hydrothermal vent sites from the East Pacific Rise in the North and South Pacific and Galapagos Rift (open circles), and vent sites from disjunct ridges in the northeast Pacific at 41–49° N (crossed circles).

type, at similar latitude and depth, and only ~1,000 kilometres apart. This distribution pattern is consistent with Johannesson's hypothesis that many taxa adapted to seamount conditions limit their dispersal to maintain their populations, because of the generally small size of seamounts, the considerable distance between them and their unique oceanographic environment^{12,13}. Circulation cells in the vicinity of seamounts and along seamount chains have been shown to facilitate recruitment to seamounts within relatively close proximity of each other and along ridge systems^{14,15}.

Endemism of the benthic fauna on southwest Pacific seamount groups/chains appears to be greater than that of deep-sea vent communities, another specialized deep-sea habitat. For comparison, we calculated C on the basis of available data for vent sites from the Explorer, Juan de Fuca and Gorda Ridges in the northeast Pacific at 41–49° N latitude, several sites along the northern East Pacific Rise (9–13°, 21° and 27° N latitude), the Galapagos Rift on the equator and 17–18° S latitude on the southern East Pacific Rise¹⁶. In contrast to the seamounts, there was substantial similarity in community composition, even at distances of 3,000–4,000 km (Fig. 3). There was some indication that similarity might be reduced between disjunct ridge systems (for example, between the Juan de Fuca and East Pacific Rise) but these differences were not significant (analysis of covariance (ANCOVA), $F_{1,12} = 3.02$, $P = 0.11$), and the similarity of vent communities between disjunct ridge systems was still clearly greater than between seamounts in the northern and southern Tasman Sea (Fig. 3). Vents are relatively ephemeral environments, whereas seamounts are typically millions to tens of millions of years in age, so dispersal between vent sites is essential to the continuance of these communities. Cold seeps, on the other hand, appear to have higher levels of endemism: most species are apparently restricted to single regions¹⁷, which has been related to the stability and isolation of these habitats. However, direct comparison with cold seep communities would be premature, as most cold seeps are poorly explored with only a few species recorded per site.

Wilson and Kaufmann³ estimated the endemism of fishes and invertebrates on seamounts as 12 and 15%, respectively, but half of the reports that they cited from reasonably well-studied regions (that is, the Hawaiian Ridge, Mid-Pacific Mountains, Kyushu-Palau Ridge, Vema seamount) indicated levels of endemism of 22–36%, consistent with our results. Citing Hubbs's 'stepping-stone' hypothesis⁵, Wilson and Kaufmann³ emphasized the faunistic links among seamount chains extending around the rim of the North Pacific, along the East Pacific Rise into the southeast Pacific and through the Emperor Chain and Hawaiian Ridge into the central North Pacific. The apparently distinct seamount chains of the southwest Pacific, Indian and Southern Oceans and much of the Indo-West Pacific¹⁸ were, however, virtually unexplored at that

time. Seventeen new genera were obtained from the Norfolk Ridge samples, four from the Lord Howe Ridge and seven or eight from the Tasmanian samples. Some appear to be relicts of groups earlier believed to have disappeared in the Mesozoic^{19–23}. These seamounts thus appear to be isolated marine systems and provide an exceptional opportunity to examine evolution and speciation in the deep sea.

The highly localized distribution of many seamount species has profound implications for their conservation. Seamount communities are extremely vulnerable to the impacts of fishing: their limited fixed habitat, the extreme longevity of many species (of the order of 100 years and more^{22–24}), and the apparently limited recruitment between seamounts, all compound the uncertainty of recovery from trawling, which sweeps away the benthic epifaunal community as a bycatch⁷. The global status of seamount benthic communities is unknown; however, the localized distribution of many benthic seamount species greatly increases the threat of extinction and may require that conservation and protection of seamount communities be undertaken on a local scale. □

Methods

Our study is based on 23 cruises that the Muséum National d'Histoire naturelle and the Institut Français de Recherche Scientifique pour le Développement en Coopération (MUSORSTOM) carried out since 1984 to study the fauna of the seamounts, ridges and adjacent seafloor within the EEZ of New Caledonia at the northern edge of the Tasman Sea, 11 of which focused on seamounts, primarily along the Norfolk and Lord Howe Ridges¹⁰; a single cruise carried out by the Commonwealth Scientific and Industrial Research Organization (CSIRO) in 1998 to study the benthic fauna of 14 small seamounts off southern Tasmania at the southern extreme of the Tasman Sea⁷; and surveys by the Museum of Victoria of the soft-bottom fauna of the continental slope off southeastern Australia based cruises in 1984, 1986 and 1994 (ref. 9) (Table 1, Fig. 1). ORSTOM carried out 462 trawls and 748 dredge hauls using a cod end with 10 mm mesh²²; many results are presented in the 21 volumes of *Resultats des campagnes MUSORSTOM*²⁶. On the CSIRO cruise, so far as was possible, each seamount was sampled on its summit, slope and base using a dredge with a cod end of 25 mm of stretched mesh; 34 samples were obtained in all.

To compare the similarity in fauna between two seamounts, we used the coefficient of community (variously referred to as Sorensen's, Dice's or the Bray–Curtis similarity statistic for binary data): $C = 2a/(b+c)$, where a is the number of species in common to both sites and b and c are the total number of species found at each site. C is commonly assumed to vary between 0 and 1 (ref. 25), but in fact the maximum value is sensitive to the relative number of species present at the two sites²⁷ and is conservatively biased when the sampling and number of species at two sites differs significantly. The proportion of species at the less well-sampled site in common to the two sites (that is, $C' = a/b$, where $b < c$) is biased in the opposite direction and was used as an upper limit to the coefficient of community between two sites, where there were large differences of sampling intensity; however, use of C' did not materially affect any of our conclusions.

Received 22 November 1999; accepted 28 March 2000

- Roden, G. I. In *Environment and Resources of Seamounts in the North Pacific* (eds Uchida, R. N., Hayasi, S. & Boehlert, G. W.) NOAA Technical Report NMFS 43, 3–12 (US Dept. of Commerce, Seattle, 1986).
- Genin, A., Dayton, P. K., Lonsdale, P. F. & Spiess, F. N. Corals on seamounts provide evidence of current acceleration over deep sea topography. *Nature* **322**, 59–61 (1986).
- Wilson, R. R. & Kaufmann, R. S. Seamount biota and biogeography. *Geophys. Monographs* **43**, 355–377 (1987).
- Rogers, A. D. The biology of seamounts. *Adv. Mar. Biol.* **30**, 305–350 (1994).
- Hubbs, C. L. Initial discoveries of fish faunas on seamounts and offshore banks in the eastern Pacific. *Pacif. Sci.* **13**, 311–316 (1959).
- Smith, D. K. & Jordon, T. H. Seamount statistics in the Pacific Ocean. *J. Geophys. Res.* **93**, 2899–2919 (1988).
- Koslow, J. A. et al. The seamount benthic macrofauna off southern Tasmania: community structure and impacts of trawling. *Mar. Ecol. Prog. Ser.* (submitted).
- Koslow, J. A., Bulman, C. M. & Lyle, J. M. The mid-slope demersal fish community off southeastern Australia. *Deep-Sea Res.* **41**, 113–141 (1994).
- Poore, G. C. B., Just, J., & Cohen, B. F. Composition and diversity of Crustacea Isopoda of the southeastern Australian continental slope. *Deep-Sea Res.* **41**, 677–693 (1994).
- Richer de Forges, B. Explorations for bathyal fauna in the New Caledonian economic zone. *Mem. Mus. Natl. Hist. Nat.* **145**, 9–54 (1990).
- Vaughan, T. W. & Wells, J. W. Revision of the suborders, families, and genera of the Scleractinia. *Geol. Soc. Am. Special Papers* **44**, 363 (1943).
- Johannesson, K. The paradox of Rockall: why is a brooding gastropod (*Littorina saxatilis*) more widespread than one having a planktonic larval dispersal stage (*L. littorea*)? *Mar. Biol.* **99**, 507–513 (1988).
- Parker, T. & Tunnicliffe, V. Dispersal strategies of the biota on an oceanic seamount: implications for ecology and biogeography. *Biol. Bull.* **187**, 336–345 (1994).
- Lutjeharms, J. R. E. & Heydorn, A. E. F. The rock lobster *Jaanus tristiani* on Vema seamount: drifting buoys suggest a possible recruiting mechanism. *Deep-Sea Res.* **28**, 631–636 (1981).
- Mullineaux, L. S. & Mills, S. W. A test of the larval retention hypothesis in seamount-generated flows. *Deep-Sea Res.* **44**, 745–770 (1997).

16. Tunnicliffe, V., McArthur, A. G. & McHugh, D. A biogeographical perspective of the deep-sea hydrothermal vent fauna. *Adv. Mar. Biol.* **34**, 353–442 (1998).
17. Sibuet, M. & Olu, K. Biogeography, biodiversity and fluid dependence of deep-sea cold-seep communities at active and passive margins. *Deep-Sea Res. II* **45**, 517–567 (1998).
18. Heezen, B. C. & Tharp, M. *World Ocean Floor* (Map). United States Navy, Office of Naval Research (Washington, D.C., 1977).
19. Bouchet, P. & Metivier, B. Living Pleurotomariidae (Mollusca: Gastropoda) from the South Pacific. *New Zealand J. Zool.* **9**, 309–318 (1982).
20. Ameziane-Cominardi, N., Bourseau, J. P. & Roux, M. Les crinoïdes pèdoncules de Nouvelle-Calédonie (S.W. Pacifique) : une faune ancestrale issue de la Mesogée mésozoïque. *C.R. Acad. Sci. Paris*, **304**, 15–18 (1987).
21. Laurin, B. Découverte d'un squelette de soutien du lophophore de type "crura" chez un brachiopode inarticulé : description de *Neoancistrocrania norfolki* gen. sp. nov. (Cranidae). *C.R. Acad. Sci. Paris*, **314**, 343–350 (1992).
22. Richer de Forges, B. La Diversité du Benthos Marin de Nouvelle-Calédonie: de l'Espèce à la Notion de Patrimoine. PhD Thesis, Muséum National d'Histoire Naturelle, Paris (1998).
23. Vacelet, J. et al. A colonial sphinctozoan sponge related to Triassic reef builders surviving in deep water off New Caledonia. *C.R. Acad. Sci.* **314**, 379–385 (1992).
24. Grigg, R. W. Resource management of precious corals: a review and application to shallow water reef-building corals. *Marine Ecol.* **5**, 57–74 (1984).
25. Pielou, E. C. in *Population and Community Ecology* 311–312 (Gordon and Breach Science, New York, 1974).
26. Richer de Forges, B. in *Resultats des Campagnes MUSORSTOM*. Volume 10 (ed. Crosnier, A.) *Mem. Mus. Natl. Hist. Nat.* **156**, 475–491 (1993).
27. Koslow, J. A. Community structure in North Atlantic deep-sea fishes. *Prog. Oceanogr.* **31**, 321–338 (1993).

Acknowledgements

Results from the northern Tasman Sea were obtained as part of the MUSORSTOM program, a collaboration between ORSTOM and the Natural History Museum in Paris, and involved some 181 researchers from 92 Institutions and 24 nations. Work around Tasmania was supported by CSIRO and grants from Environment Australia and the Fisheries Research Development Corporation and involved 25 researchers. Their contributions are gratefully acknowledged; in particular, we thank R. Grandperrin, A. Crosnier, P. Bouchet, K. Gowlett-Holmes, M. Lewis, S. Cummings, J.S. Philippe and the crews of the RV *Alis* and *Southern Surveyor*. J.A.K. received support from a National Research Council Senior Research Fellowship.

Correspondence and requests for materials should be addressed to J.A.K. (e-mail: tony.koslow@marine.csiro.au).

Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit

Richard H. R. Hahnloser^{†‡}, Rahul Sarpeshkar^{†§}, Misha A. Mahowald^{*}, Rodney J. Douglas^{*} & H. Sebastian Seung^{†‡}

^{*} Institute of Neuroinformatics ETHZ/UNIZ, Winterthurerstrasse 190, 8057 Zürich, Switzerland

[†] Bell Laboratories, Murray Hill, New Jersey 07974, USA

[‡] Department of Brain and Cognitive Sciences and [§] Department of Electrical Engineering and Computer Science, MIT, Cambridge, Massachusetts 02139, USA

Digital circuits such as the flip-flop use feedback to achieve multistability and nonlinearity to restore signals to logical levels, for example 0 and 1. Analogue feedback circuits are generally designed to operate linearly, so that signals are over a range, and the response is unique. By contrast, the response of cortical circuits to sensory stimulation can be both multistable and graded^{1–4}. We propose that the neocortex combines digital selection of an active set of neurons with analogue response by dynamically varying the positive feedback inherent in its recurrent connections. Strong positive feedback causes differential instabilities that drive the selection of a set of active neurons under the constraints embedded in the synaptic weights. Once selected, the active neurons generate weaker, stable feedback that provides analogue amplification of the input. Here we present our

model of cortical processing as an electronic circuit that emulates this hybrid operation, and so is able to perform computations that are similar to stimulus selection, gain modulation and spatio-temporal pattern generation in the neocortex.

The multistability of digital circuits and the linear amplification achieved in analogue circuits are generally seen as incompatible functions and are separated into two classes of electronic technology. However, the neuronal circuits of the neocortex do not respect this distinction. There, multistability coexists with analogue response. For example, when a visual stimulus is attended at the expense of other visual stimuli—the subject is concentrating on one object in a field of vision—then many cortical neurons tend to respond in a graded way to the sensory attributes of the attended stimulus, as if it were presented alone².

We have designed a simple electronic circuit that emulates this hybrid behaviour. The circuit comprises a ring of 16 excitatory neurons, each of which makes synaptic connections of variable strength onto itself and onto its nearest and next nearest neighbours. These localized excitatory interactions reflect the preference for local connections observed in neocortex. At the centre of the ring is a single inhibitory neuron that receives synaptic input from all the

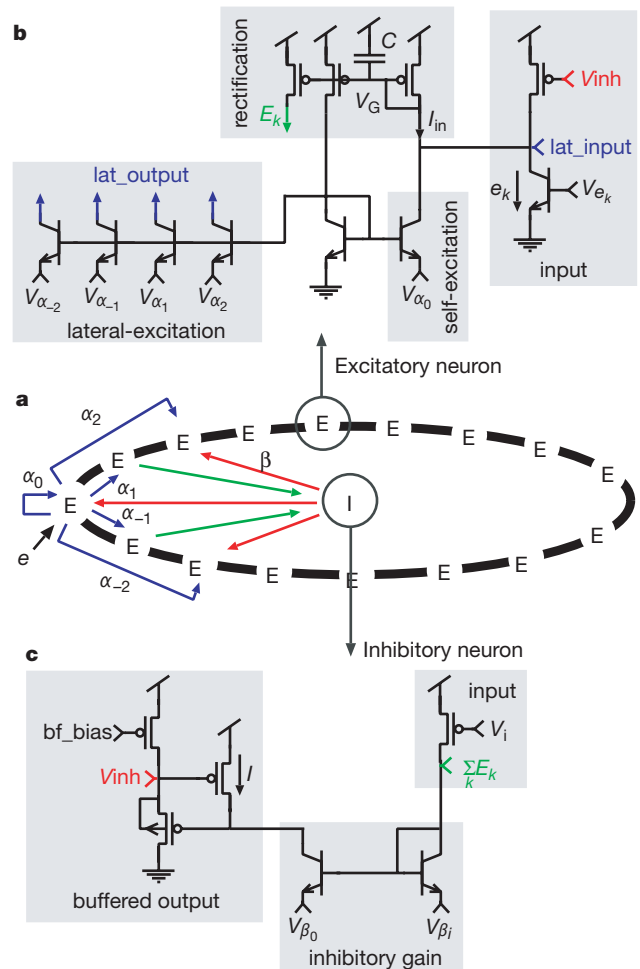


Figure 1 Silicon implementation of a recurrent network with a ring architecture. The synaptic connections are consistently coloured in the a–c. **a**, Each excitatory neuron E makes connections (blue) with itself and four neighbours on the ring, of strength α_i . The inhibitory neuron I makes global recurrent connections on the ring of strength β . Green, presynaptic connections; red, postsynaptic connections. **b**, The core of the excitatory neuron is the current mirror labelled rectification which contains a large capacitance. **c**, The core of the inhibitory neuron is the bipolar current mirror that sets the gain of inhibition (see Methods).

excitatory neurons, and returns inhibition to them. This simple architecture and similar variants have been used previously to model response properties of neurons in cortex^{5–9} and other^{10–12} brain areas.

The output of each excitatory neuron is an electrical current that is positive if the neuron is active, and zero if it is inactive. Negative values are not possible, because the artificial neurons are based on current mirrors, which have a rectification nonlinearity (Fig. 1b, rectification).

Each excitatory neuron can be stimulated independently by an electrical current. The response of the population to stimulation of a single neuron is shown in Fig. 2a (red line). The response profile is centred on the stimulus and extends over a large fraction of the ring. In this way, the output currents of the circuit form a distributed representation of stimulus location. In many brain areas⁸, such distributed representations have been observed and been referred to as population codes^{13,14}. In our circuit, the population code arises by recurrent excitation, which causes spreading of activity in both directions, until it is cut off by recurrent inhibition. Beyond this cut-off, the excitation is not strong enough to overcome the threshold for activation set by inhibition.

To generate Fig. 2a, the stimulus was fixed, and the responses of all the neurons were measured. A similar graph is obtained if the response of a particular neuron is plotted as a function of the location of the stimulus, which is the traditional procedure used by electrophysiologists to map a receptive field or the tuning curve of a cell. Neurons can maintain their tuning to a sensory variable, such as stimulus orientation or retinal location, but respond with an amplitude that is modulated by another variable, such as eye position or attention. Our circuit exhibits this remarkable phenomenon, called gain modulation, when a uniform background excitation is applied in addition to the fixed localized stimulus.

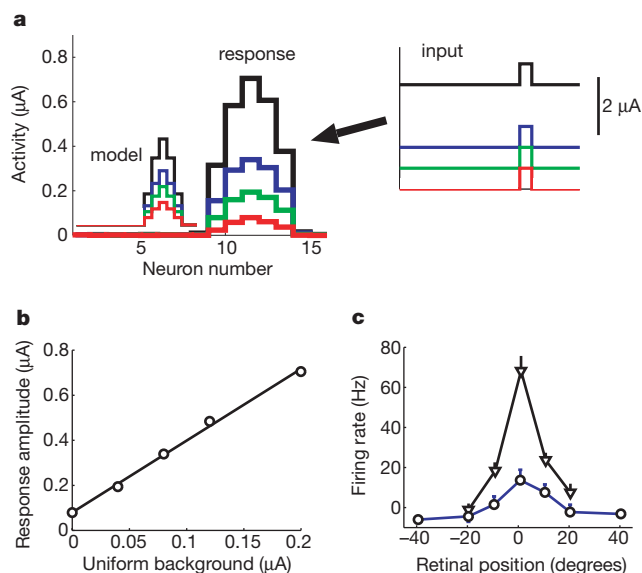


Figure 2 Modulation of tuned population response. $V_{\alpha-1} = 5 \text{ mV}$, $V_{\alpha_1} = 6 \text{ mV}$, $V_{\alpha-2} = 5 \text{ mV}$, $V_{\alpha_2} = 9 \text{ mV}$, $V_{\alpha_0} = 88 \text{ mV}$, $V_{\beta_1} = 11 \text{ mV}$ and $V_{\beta_0} = 6 \text{ mV}$. **a**, The thick red line represents the response of the circuit to a single stimulus applied to neuron number 12. When a uniform background input is added, the response profile is modulated but its shape remains approximately invariant. The colours indicate which output profile corresponds to which input profile shown on the right. **b**, The response of neuron 12 in **a** is plotted versus the amplitude of the uniform background, resulting in a linear relationship. **c**, Gain modulation as observed in posterior parietal cortex (reproduced with permission from R. Andersen; ref. 15). Firing rates for two different eye positions are plotted as a function of the retinal horizontal eccentricity of a visual stimulus. Each data point is the average response to eight repetitions of the stimulus. Background activity before the stimulus presentation has been subtracted from the total activity during measurement.

When we changed the amplitude of the background, the population response remained at the same location with much the same shape in Fig. 2a, but with an amplitude that varied with background amplitude in an approximately linear way (Fig. 2b). Thus, the background modulated the amplitude of the tuning curve of each neuron. For comparison, an example of gain modulation observed in posterior parietal cortex¹⁵ is shown in Fig. 2c. The tuning curve indicates that the neuron is selective for the location of a visual stimulus in retinotopic coordinates, while the amplitude of response linearly encodes the position of the eyes. In a previous model of gain modulation, it was assumed that excitatorily coupled neurons in parietal cortex share the same efference copy of eye position, which amounts to a uniform background input⁷.

Whereas the foreground stimulus and the background cooperate to determine the response of the circuit in Fig. 2, two localized stimuli compete to determine the response in Fig. 3. The circuit selects one of the stimuli while completely suppressing its response to the other (Fig. 3a, b). When the amplitudes of the two stimuli are sufficiently different from each other, the circuit always selects the larger stimulus. This property is evident in the hysteresis curve (Fig. 3c), which has a single branch at the extremes. By contrast, when the stimuli have roughly equal amplitudes, the circuit may select either one stimulus or the other. In this bistable regime, once a choice has been made, the circuit maintains its past selection and is insensitive to small changes in relative amplitudes of the stimuli. Bistability enables the neurons to maintain a memory of the past, but in spite of bistability, the neural firing rates remain continuously graded. For example, if the amplitudes of the stimuli are rescaled by the same factor, then the response amplitude is similarly rescaled. On the other hand, in Fig. 3d, when the stimuli marked by asterisks are moved closer to each other, then the circuit tends to interpolate between them rather than to select one of them. The network

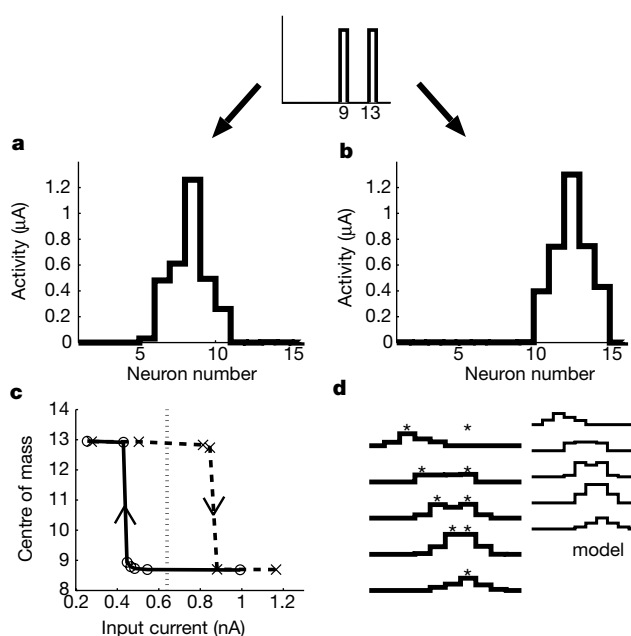


Figure 3 Multistability of selection. Synaptic strengths are the same as in Fig. 2. In response to two identical stimuli (inset, $V_{\alpha_1} = V_{\alpha_2} = 0.425 \text{ V}$), the circuit selects one and suppresses the other. The left stimulus is selected in **a** and the right stimulus is selected in **b**. **c**, Hysteresis is evident in a graph of the response centroid. The intensity of the left stimulus (referred to as input current on the x-axis) is swept upward (dashed line) and then downward (solid line), whilst the intensity of the right stimulus is held fixed (dotted vertical line). **d**, Averaging versus selection. When two identical stimuli (asterisks) are nearby, then excitatory neurons in between the stimuli respond to their presence (lower traces). When the two stimuli are further apart, then the circuit selects one of the stimuli (here, the left stimulus, top trace).

interprets two nearby stimuli as a corrupted version of a single intermediate one, and restores its output towards the choice.

In trying to understand the analogue–digital nature of our specific circuit design, we have developed a general theory of recurrent networks with rectification nonlinearity and symmetric interactions. The starting point is a basic observation about the strength of feedback. Even if all synaptic strengths are kept fixed, the effective strength of feedback is variable. This property arises because the neurons of the network can be partitioned into two disjoint sets: those that are above threshold and active, and those that are below threshold and inactive. The effective strength of

feedback depends only on the synaptic connections between active neurons, as only they participate in feedback loops. Depending on this effective strength, each active set can be classified as either ‘forbidden’ or ‘permitted’. Forbidden sets are those for which feedback is strong enough to induce instability. A forbidden set cannot be realized by a stable steady state, no matter what external input is applied. In contrast, permitted sets are those for which feedback is weak enough that there are no instabilities. They can be realized by a stable steady state, given the proper pattern of external input. The fact that the stability of a steady state depends only on the identities of active neurons, and not on their analogue responses, is

Box 1

Theory of symmetric networks with rectification

In general, our current-mode design techniques can be used to implement electronic circuits that are approximately described by the dynamical equations

$$\tau_i(x_i) \frac{dx_i}{dt} + x_i = \left[b_i + \sum_j W_{ij} x_j \right]_+ \quad (1)$$

Here the variables $\mathbf{x} = (x_i) \geq 0$ denote the output currents of neurons and $\mathbf{b} = (b_i)$ their input currents. The strength of the synaptic connection from the j th to the i th neuron is denoted by W_{ij} and $[x]_+ = \max(0, x)$ is a rectification nonlinearity. The time scales $\tau_i(x_i) \propto (x_i)^{-1}$ are inversely proportional to the x_i , which is a basic property of current-mode circuits. For the proofs of the results to follow, see Supplementary Information.

Effective gain at steady state

We assume that the inputs $\mathbf{b}$ are constant in time, and analyse the steady states $\bar{x}_i$ defined by $\frac{dx_i}{dt} = 0$. In such steady states, digital selection and analogue amplification can be separated from each other by using binary variables σ_i taking the values 1 or 0, depending on whether the i th neuron is active ($\bar{x}_i > 0$) or inactive ($\bar{x}_i = 0$).

$$\bar{\mathbf{x}} = G\mathbf{b}, \quad \text{where } G = (I - \Sigma W \Sigma)^{-1} \Sigma \quad (2)$$

is called the effective gain matrix²⁵. I denotes the identity matrix and $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_N)$ is a matrix that carries ‘digital’ information about which neurons are active ($\sigma_i = 1$ if neuron i is active and zero otherwise). The effective gain G depends on the identities of the active neurons, though it is independent of their analogue responses. It is composed of two factors. One factor $(I - \Sigma W \Sigma)^{-1}$ quantifies the feedback amplification mediated by the synaptic connections between active neurons (the expression $\Sigma W \Sigma$ is like W , except that the strengths of all synapses involving inactive neurons are zeroed out). The other factor, Σ , signifies that only the inputs to active neurons are amplified.

Global stability if no unstable common mode

Do the dynamics in equation (1) always converge to a steady state? In the case of some symmetric networks ($W_{ij} = W_{ji}$), the answer is yes. Global stability can be proven if $\Sigma W \Sigma$ has no unstable common modes for any Σ , where a common mode is an eigenvector whose components all have the same sign and instability means that the corresponding eigenvalue is larger than one. Intuitively, unstable common modes lead to runaway excitation, incompatible with global stability.

On the other hand, a differential mode of $\Sigma W \Sigma$ is an eigenvector with both positive and negative components. In contrast to common modes, a differential mode can be unstable and not cause any harm to the global stability of a network. The growth of unstable differential modes necessarily makes some neurons hit rectification nonlinearity. As we will see, this turns out to be beneficial for stability.

Local stability is independent of analogue response values

The local stability of steady states can be determined by linearizing the dynamics about $\bar{\mathbf{x}}$,

$$\frac{d\mathbf{x}}{dt} = -D(\bar{\mathbf{x}})^{-1} (I - \Sigma W \Sigma) (\mathbf{x} - \bar{\mathbf{x}}) \quad (3)$$

where $D(\bar{\mathbf{x}}) = \text{diag} \tau_1(\bar{x}_1), \dots, \tau_N(\bar{x}_N)$. Using the assumption of a symmetric W , local stability follows if the largest eigenvalue of $\Sigma W \Sigma$ is bounded above by unity (it can be proven that the factor $D(\bar{\mathbf{x}})^{-1}$ does not matter). This means that the stability of a steady state depends on the set of active neurons, but does not depend on their analogue responses. In practice, only stable steady states can be observed in a network, because even infinitesimal amounts of noise cause divergence from an unstable steady state. However, differential instabilities are essential for digital constraints on selection, because they are necessary for the existence of forbidden sets.

The grouping into permitted and forbidden sets

If the largest eigenvalue of $\Sigma W \Sigma$ is greater than unity for some set of active neurons defined by Σ , we refer to that set as a ‘forbidden set’. On the other hand, if the largest eigenvalue is bounded above by unity, the active set is ‘permitted’. It can be shown that for every permitted set, there exists a stimulus vector $\mathbf{b}$ that produces a stable steady state response consistent with that set.

Our main result is that in a symmetric network, any subset of a permitted set is permitted and any superset of a forbidden set is forbidden. Proof: this grouping follows from the fact that the largest eigenvalue of $\Sigma W \Sigma$ is always smaller than that of W . The largest eigenvalue of W can be characterized by using the Rayleigh–Ritz quotient²⁶:

$$\lambda_{\max} = \max_{\mathbf{v}} \frac{\mathbf{v}^T W \mathbf{v}}{\mathbf{v}^T \mathbf{v}} \quad (4)$$

Similarly, the largest eigenvalue of $\Sigma W \Sigma$ is given by maximizing the same expression, but constraining $v_i = 0$ for the inactive neurons. As imposing more constraints cannot increase the maximum, it follows that inactivating a neuron cannot increase the strength of positive feedback and therefore cannot enhance instability. Intuitively, both mutual excitation and mutual inhibition constitute positive feedback and so explain this hierarchical grouping of permitted sets.

For the mathematical model of our circuit with the parameters given in the Methods section, any active set consisting of more than five contiguous neurons is a forbidden set, and so sets a maximal response width (this result was derived by numerically diagonalizing $\Sigma W \Sigma$).

No multistable selection without differential instabilities

The circuit illustrated in Figs 2 and 3 is potentially multistable, in the sense that there exists at least one input vector $\mathbf{b}$ allowing for more than a single stable steady state. In general, the existence of differential instabilities is necessary and sufficient for potential multistability. This can be proven using Lyapunov function arguments. Intuitively, multistability occurs when there is more than one local minimum of the Lyapunov function. In this situation, saddles exist between the multiple minima, indicating that there are unstable differential modes.

a specific consequence of rectification nonlinearity (Box 1).

In the course of converging to a steady state, the dynamics of a network of N neurons must select one of the 2^N possible active sets. The existence of forbidden sets gives the selection process a digital character, meaning that the active sets at steady state obey hard constraints. Most importantly, the constraints cannot be violated, no matter what the external input may be. In general, varying the recurrent synaptic strengths of a network causes these constraints to change. In particular, if synaptic feedback is weak, then there may be no forbidden sets at all. Similarly, networks with different patterns of synaptic connectivity generally implement different constraints. However, this flexibility has its limits, as the constraints cannot take an arbitrary form. It can be shown that any subset of a permitted set is permitted, and any superset of a forbidden set is forbidden (Box 1).

Given that the network dynamics has selected a permitted set, the analogue outputs of the active neurons are related to their inputs by an effective gain matrix, which quantifies the amplification or attenuation produced by the feedback connections between the active neurons. As explained earlier, inactive neurons and their synaptic connections are irrelevant. The effective gain depends on the identities of the active neurons only, though it does not depend on their analogue responses. The result is that steady states with the same active set behave like the responses of a linear system, but nonlinearities emerge when different active sets are compared with each other.

This general theory is applicable to our circuit in Figs 2 and 3, where there is approximate symmetry of the synaptic interactions (in the case of nonsymmetric interactions, the theory does not apply because there can be spontaneous generation of spatiotemporal activity patterns; see Supplementary Information). Due to local cooperation and lateral competition, the number of contiguous neurons in a permitted set is limited and so activity profiles cannot exceed a maximal width. This does not mean that there is a similar constraint that prevents the narrowing of an activity profile. Narrowing can be achieved, for example, by applying a feedforward input i (Fig. 1) to the inhibitory neuron. This property follows from our earlier statement that every subset of a permitted set is also a permitted set and is the reason why the response does not broaden in Fig. 2, even when the input becomes almost completely uniform^{5,7,16}. Intuitively, a uniform, nonspecific input transiently excites the most unstable mode in any symmetric network and so is a computational means of enforcing selection of a permitted set comprising many active neurons. The suppression of the response to one stimulus in Fig. 3a and b can also be understood as the outcome of constrained selection. Constraints also cause the circuit to interpret two nearby stimuli as a corrupted version of a single, intermediate stimulus (Fig. 3d).

The constraints that define the maximal response width to stimulation can be made more or less restricted by changing the synaptic strengths. For example, if the strengths of lateral excitation are adjusted to zero, leaving only self-excitation, then the only permitted sets are those with a single neuron active, so that the circuit functions in a winner-take-all mode^{17–19}.

The concept of effective gain is useful for understanding the analogue behaviours observed in Figs 2 and 3. In Fig. 2a, the set of active neurons remains basically the same for all stimuli, so that the effective gain is constant, explaining the linearity of Fig. 2b. By contrast, the active sets are completely different in Fig. 3a and b, and so are the effective gains, which explains why the response is so nonlinear.

The multistability shown in Fig. 3 is reminiscent of the flip-flop, a digital feedback circuit consisting of two elements interacting by mutual inhibition. The feedback creates differential instability, which drives the voltages of the two elements to opposite limits of the power supply. This is why the flip-flop is constrained to two possible states, and is bistable. Similarly, in our general theory of recurrent networks, the existence of differential instabilities is

necessary for constrained selection, and the possibility of multistability (Box 1). But these networks differ from the flip-flop, in that growth of instabilities is limited by threshold nonlinearity alone. No saturation or upper limit on activity is necessary to hold instabilities in check. That is why digital selection can coexist with analogue response.

Our theory of recurrent networks is related to, but distinct from, the paradigm of computation with dynamical attractors^{20,27}. According to the attractor model, memories are stored as stable fixed points, or dynamical attractors of a recurrent neural network. External input serves only to initialize the network dynamics, which retrieves a stored memory by converging to an attractor. The attractor model was exciting because it provided a mathematically precise formulation of the intuitive idea that patterns can be latent in the recurrent connections of a network. Our theory provides a different formalization of the same intuitive idea. The permitted sets can be regarded as binary patterns stored in the recurrent connections. The identities of active neurons are determined by retrieving one of these stored memories. However, this retrieval does not fix the analogue responses of the active neurons. These arise from feedback amplification and attenuation of the maintained external input.

In our circuit, the permitted sets are simply groups of neighbouring neurons. However, permitted sets can take more complex forms in circuits with more sophisticated synaptic connectivity. For example, we can imagine a computational role for permitted sets in a network for visual recognition of letters of the alphabet. This network would contain neurons that represent the strokes of which letters are composed. Then co-activation of the neurons in a permitted set would represent a letter as a combination of strokes. Other combinations that do not correspond to letters would be forbidden by the recurrent connections. More generally, in parts-based representations of objects²¹, forbidden sets could impose 'syntactic' constraints on the way that parts are combined to form a whole. But analogue variability is also an important property of perceptual stimuli. Changes in viewpoint, lighting, and other factors can cause continuous variations in the image of an object. As both analogue computation and digital constraints coexist in our circuits, they are potentially well suited for applications to machine perception and efficient hybrid computations²². To build such circuits, it will be important to devise mechanisms of synaptic plasticity, so that parts-based representations for perception can be learned. □

Methods

Our circuit was fabricated using the 2 μm process of the MOSIS facility. Bipolar transistors were used, as well as MOSFETS (metal oxide semiconductor field effect transistors) because in weak inversion current mirrors made from bipolar transistors have superior matching and a larger linear range than those made from MOSFETS. We have used a current-mode design, which is a suitable technique for implementations of neural circuits²³. In addition to the ring network of Fig. 1, the chip included digital circuitry for scanning the multiple currents of the excitatory neurons²⁴ out through a single current-sense amplifier.

We can derive model equations for the currents E_k of excitatory neurons and I of the inhibitory neurons (see Fig. 1). By substituting the current–voltage relationship $E_k = I_0 e^{qV_G/U_T}$ of a CMOS (complementary metal-oxide semiconductor) transistor ($\kappa \approx 0.7$ is the gate efficiency constant, $U_T \approx 25$ mV is the thermal voltage and I_0 is the dark current) into $I_m - E_k = C \frac{dV_G}{dt}$ for the gate voltage V_G of the rectification mirror (here V_G is referenced to VDD) we obtain the output currents E_k ($k = 1, \dots, N$) of the $N = 16$ excitatory neurons:

$$C \frac{dE_k}{dt} + E_k = \left[e_k + \sum_{l=2}^2 \alpha_l E_{k+l-1} - \beta I \right]_+ \quad (5)$$

$$I = i + \sum_{k=1}^N E_k \quad (6)$$

In this derivation, $E_{N+l} = E_l$ has been defined and $C > 0$ is a constant depending on the capacitance of the rectification mirror. Our measurements indicate that equations (5) and

(6) are valid for output currents ranging over six orders of magnitude, from 1 nA to tens of mA of current.

The excitatory neurons interact by self (α_0), nearest neighbour (α_1 and α_{-1}), and next nearest neighbour (α_2 and α_{-2}) connections. The inhibitory neuron I sums the currents from the excitatory neurons, and in turn inhibits them with a strength β . The strength α_i of each type of excitatory synapse is globally controlled by the emitter voltage V_{α_i} and the strength β of inhibition is globally controlled by the difference $V_{\beta_i} - V_{\beta_0}$: $\alpha_i = e^{-V_{\alpha_i}/U_T}$ and $\beta = e^{(V_{\beta_i} - V_{\beta_0})/U_T}$ (the emitter voltages are references to a variable cell ground). Similarly, the input currents e_k and i to excitatory and inhibitory neurons are exponentially related to externally controlled voltages V_{e_k} and V_i . A synaptic weight can be turned off by setting the emitter voltage to about 1V, where it is above the base voltage, under which conditions current cannot flow across the bipolar.

In excitatory-inhibitory networks, it is known that slow inhibition can lead to oscillatory or explosive instability. Therefore we sped up inhibition by buffering the current mirrors of the inhibitory synapses ('buffered output' in Fig. 1c, $\text{bf_bias} = 3.5$ V), and we slowed down excitation by adding a large capacitance to the current mirror of the excitatory neurons. In equation (6), the speed of inhibition was approximated as instantaneous.

We were able to reproduce the measurements shown in Figs 2 and 3 by simulating the model equations (5) and (6) with parameter values $\alpha_0 = 0$, $\alpha_{-1} = \alpha_1 = 1.15$, $\alpha_{-2} = \alpha_2 = 8$ and $\beta = 0.5$. The parameters of the model were tuned to match the response profiles. Although we tried to minimize the mismatch between different neurons and synapses by a generous layout in terms of transistor and capacitor size, there are small inhomogeneities owing to imperfections of the fabrication process.

Received 27 December 1999; accepted 13 April 2000.

1. Leopold, D. A. & Logothetis, N. K. Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* **379**, 549–553 (1996).
2. Treue, S. & Maunsell, J. H. Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* **382**, 539–541 (1996).
3. Treue, S., Martinez, S. & Trujillo, J. C. Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* **399**, 575–579 (1999).
4. Reynolds, J. H., Chelazzi, L. & Desimone, R. Competitive mechanisms subserve attention in macaque areas V2 and V4. *J. Neurosci.* **19**, 1736–1753 (1999).
5. Ben-Yishai, R., Lev Bar-Or, R. & Sompolinsky, H. Theory of orientation tuning in visual cortex. *Proc. Natl. Acad. Sci. USA* **92**, 3844–3848 (1995).
6. Hansel, D. & Sompolinsky, H. In *Methods in Neuronal Modeling* (eds Koch, C. & Segev, I.) 499–567 (MIT Press, Cambridge, Massachusetts, 1998).
7. Salinas, E. & Abbott, L. F. A model of multiplicative neural responses in parietal cortex. *Proc. Natl. Acad. Sci. USA* **93**, 11956–11961 (1996).
8. Georgopoulos, A. P., Taira, M. & Lukashin, A. Cognitive neurophysiology of the motor cortex. *Science* **260**, 47–52 (1993).
9. Camperi, M. & Wang, X. J. A model of visuospatial working memory in prefrontal cortex: recurrent network and cellular bistability. *J. Comput. Neurosci.* **5**, 383–405 (1998).
10. Zhang, K. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory. *J. Neurosci.* **16**, 2112–2126 (1996).
11. Kopeck, K. & Schoner, G. Saccadic motor planning by integrating visual information and pre-information on neural dynamical fields. *Biol. Cybern.* **73**, 49–60 (1995).
12. Amari, S.-I. & Arbib, M. A. in *Systems Neuroscience* (ed. Metzler, J.) 119–165 (Academic Press, New York, 1977).
13. Seung, H. S. & Sompolinsky, H. Simple models for reading neuronal population codes. *Proc. Natl. Acad. Sci. USA* **90**, 10749–10753 (1993).
14. Abbott, L. F. Decoding neuronal firing and modelling neural networks. *Q. Rev. Biophys.* **27**, 291–331 (1994).
15. Andersen, R. A., Essick, G. E. & Siegel, R. M. Encoding of spatial location by posterior parietal neurons. *Science* **230**, 456–458 (1985).
16. Amari, S.-I. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* **27**, 77–87 (1977).
17. Lazzaro, J., Ryckebusch, S., Mahowald, M. A. & Mead, C. A. in *Advances in Neural Information Processing Systems* (ed. Touretzky, D. S.) 703–711 (Morgan Kaufmann, San Mateo, California, 1989).
18. De Weert, P. & Morris, T. G. CMOS current mode winner-take-all circuit with distributed hysteresis. *Electron. Lett.* **31**, 1051–1053 (1995).
19. Indiveri, G. Winner-take-all networks with lateral excitation. *Analogue Integrated Circuits Signal Process* **13**, 185–193 (1997).
20. Amit, D. J. The Hebbian paradigm reintegrated: Local reverberations as internal representations. *Behav. Brain Sci.* **18**, 617–657 (1995).
21. Lee, D. & Seung, H. S. Learning the parts of objects by nonnegative matrix factorization. *Nature* **401**, 788–791 (1999).
22. Sarpeshkar, R., Lyon, R. F. & Mead, C. A. low-power wide-dynamic-range analogue vlsi cochlea. *Analogue Integrated Circuits Signal Process.* **16**, 245–274 (1998).
23. Andreou, A. G. & Boahen, K. A. Synthetic neural circuits using current-domain signal representations. *Neural Computat.* **1** (1989).
24. Mead, C. & Delbrück, T. Scanners for visualizing activity of analogue VLSI circuitry. *Analogue Integrated Circuits Signal Process.* **1**, 93–106 (1991).
25. Hahnloser, R. H. R. About the piecewise analysis of networks of linear threshold neurons. *Neural Net.* **11**, 691–697 (1998).
26. Horn, R. A. & Johnson, C. R. *Matrix Analysis* (Cambridge Univ. Press, New York, 1985).
27. Hopfield, J. J. & Tank, D. W. Neural computations of decisions in optimization problems. *Biol. Cybern.* **52**, 141–152 (1985).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We acknowledge the support of the Swiss National Science Foundation SPP Program, Lucent Technologies and MIT. We thank A. Andreou, J. Kramer, D. Lee, C. Brody, M. Fee, D. Tank, P. Sinha, S. Roweis and S.-C. Liu for discussions about the circuits and current-mode design.

Correspondence and requests for materials should be addressed to R.H. (e-mail: rh@ai.mit.edu).

Induction of neurogenesis in the neocortex of adult mice

Sanjay S. Magavi, Blair R. Leavitt* & Jeffrey D. Macklis

Division of Neuroscience, Children's Hospital, and Department of Neurology and Program in Neuroscience, Harvard Medical School, Enders 350, 320 Longwood Avenue, Boston, Massachusetts 02115 USA

Neurogenesis normally only occurs in limited areas of the adult mammalian brain—the hippocampus¹, olfactory bulb^{2–4} and epithelium⁵, and at low levels in some regions of macaque cortex⁶. Here we show that endogenous neural precursors can be induced *in situ* to differentiate into mature neurons, in regions of adult mammalian neocortex that do not normally undergo any neurogenesis. This differentiation occurs in a layer- and region-specific manner, and the neurons can re-form appropriate corticothalamic connections. We induced synchronous apoptotic degeneration^{7,8} of corticothalamic neurons in layer VI of anterior cortex of adult mice and examined the fates of dividing cells within cortex, using markers for DNA replication (5-bromo-deoxyuridine; BrdU) and progressive neuronal differentiation. Newly made, BrdU-positive cells expressed NeuN, a mature neuronal marker, in regions of cortex undergoing targeted neuronal death and survived for at least 28 weeks. Subsets of BrdU+ precursors expressed Doublecortin, a protein found exclusively in migrating neurons^{9,10}, and Hu, an early neuronal marker^{11,12}. Retrograde labelling from thalamus demonstrated that BrdU+ neurons can form long-distance corticothalamic connections. Our results indicate that neuronal replacement therapies for neurodegenerative disease and CNS injury may be possible through manipulation of endogenous neural precursors *in situ*.

There is precedent for neuronal death modifying the fate of immature precursor cells, but only in regions of the vertebrate brain that have ongoing neurogenesis. Direct^{13–15} and correlative evidence^{5,16,17} suggests that neuronal death can trigger increased neuron addition in these systems. Our previous results show that in regions of adult mouse neocortex undergoing synchronous apoptotic degeneration of projection neurons^{7,8}, the surrounding cells, including interneurons, upregulate the expression of a specific set of developmental signalling molecules that may guide neocortical projection neuron differentiation¹⁸. Immature neurons or multipotent neural precursors that are transplanted into these regions migrate selectively to layers of cortex where projection neurons are degenerating^{7,8,19,20}. They then differentiate into projection neurons^{7,8,15,19,20}, receive afferent synapses^{7,15,20} and re-form appropriate long-distance connections to the original contralateral targets of the degenerating neurons^{19,20}. Similarly, induction of targeted neuronal death in projection neurons of the song circuitry in the avian forebrain causes increased neuron replacement from

* Present address: Center for Molecular Medicine and Therapeutics, Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia V5Z 4H4, Canada.

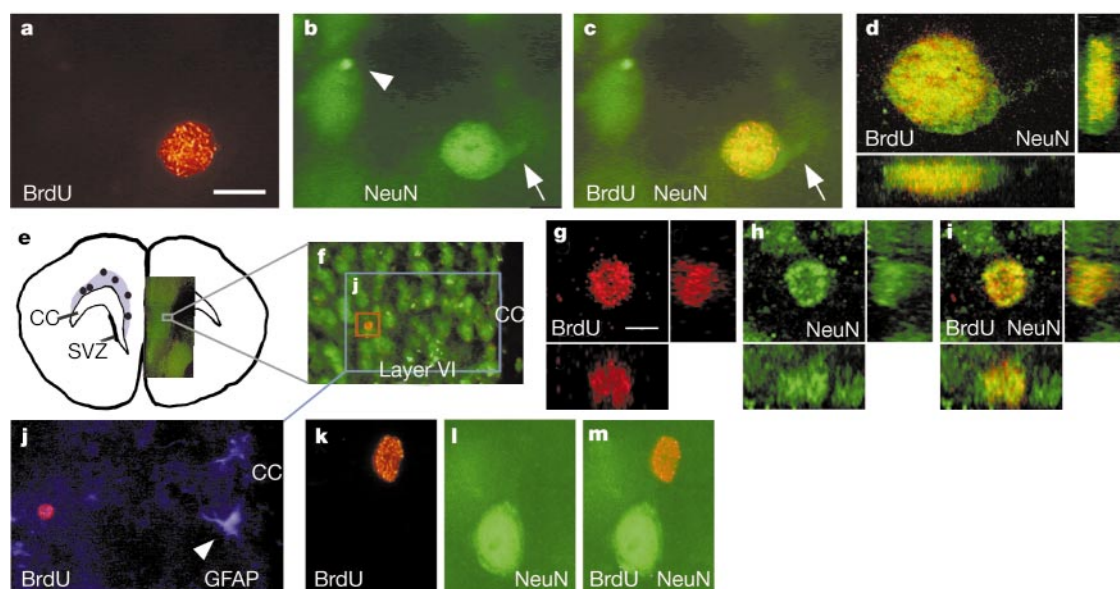


Figure 1 Newly generated BrdU+ cells can be induced to differentiate into neurons, in regions of cortex undergoing targeted apoptotic degeneration of corticothalamic neurons. **a**, A large, densely stained BrdU+ nucleus (red) 28 weeks after induction of apoptosis. Scale bar: 10 μ m. **b**, The BrdU+ neuron, lower right, shows typical staining for NeuN, a mature neuronal marker, with strong nuclear and weaker cytoplasmic labelling. Apical dendrite (arrow). The neuron at left in a different focal plane remains from the original nanosphere targeting for apoptosis; a lysosome containing nanospheres is indicated (arrowhead). No BrdU+/NeuN+ neurons contained nanospheres, demonstrating that they are not the original, targeted neurons. **c**, Overlay. **d**, Confocal images combined to produce a 3D reconstruction of this newborn neuron. Views along the x-axis (right) and y-axis (below) demonstrate co-localization of BrdU and NeuN. **e**, Left, location of BrdU+/

NeuN+ cells (dots) within the targeted region (grey); right, a sample neuron in cortical layer VI. Corpus callosum (CC). Subventricular zone (SVZ). **f**, Higher magnification view of layer VI shows a BrdU+ neuron (red box). **g–i**, Confocal 3D reconstruction of red boxed region in **f**. Right, x-axis view; bottom, y-axis view. **g**, BrdU (red), indicating that the cell underwent mitosis during the two weeks following induction of apoptosis. Scale bar: 5 μ m. **h**, NeuN (green). **i**, Merged image. BrdU (red) and NeuN (green) labelling are coincident in all three dimensions. **j**, The BrdU+/NeuN+ neuron (red) is GFAP-negative (blue). A GFAP+ astrocyte (arrowhead). **k–m**, The presence of BrdU+/NeuN– and BrdU–/NeuN+ cells demonstrate that the double labelling protocol is specific. Image from same section as **a–d**. **k**, BrdU (red) and **l**, NeuN (green) do not show crossreactivity. **m**, Overlay.

endogenous neural precursors in a system already undergoing low-level neurogenesis²¹. Endogenous multipotent neural precursors are now known to exist in the mammalian forebrain. However, numerous studies have reported that neurogenesis does not normally occur in postnatal mouse cortex, raising the question of whether this reflects an intrinsic limitation of the endogenous neural precursors' potential, or a lack of appropriate micro-environmental signals necessary for neuronal differentiation or survival.

To examine whether the developmental signals re-expressed during synchronous targeted apoptosis can direct the fate of endogenous precursors in the mature cortex, we examined the differentiation of endogenous precursors exposed *in situ* to intercellular signals modified by targeted apoptosis in adult mouse cerebral cortex, without transplantation. We induced synchronous apoptotic degeneration of layer VI corticothalamic projection neurons using chromophore-targeted neuronal degeneration. We targeted the neurons in the anterior neocortex because they are close to the subventricular zone, a region that contains a relatively large number of multipotent neural precursors. We treated mice with BrdU for two weeks and examined the phenotype of newly made, BrdU-positive (BrdU+) cells 2, 5, 9 and 28 weeks after induction of apoptosis. We examined the differentiation of these cells using markers of developing and mature neurons and glia, including Doublecortin, Hu, NeuN, glial fibrillary acidic protein (GFAP) and myelin basic protein (MBP).

To identify neurons in the adult neocortex, we used antibodies against NeuN, a transcription factor that is expressed in the nucleus and cytoplasm of mature neurons. We found BrdU+/NeuN+ neurons only in regions of cortex undergoing apoptotic degeneration of layer VI thalamocortical projection neurons at 2–28 weeks after inducing degeneration (Fig. 1). To confirm that new cells were double-labelled and not merely closely apposed to pre-existing

neurons, we imaged them using laser scanning confocal microscopy and produced three-dimensional (3D) digital reconstructions. The reconstructions confirmed that the new cells were neurons. Some newborn neurons had pyramidal neuron morphology (large, 10–15 μ m diameter somata with apical processes; Fig. 1a–d) characteristic of neurons that give rise to long-distance projections. At two weeks, we found 97 ± 69 (mean $\pm$ s.d.) new neurons per mm^3 ($n = 6$). At five weeks, there were 43 ± 20 newborn neurons per mm^3 ($n = 3$) and at 28 weeks, 78 ± 15 new neurons per mm^3 ($n = 2$). In contrast, we found no BrdU+/NeuN+ cells in the cortex of control mice at any time point ($n = 13$), confirming previous reports. The difference between control and experimental mice is highly statistically significant ($P < 0.0001$, two-tailed *t*-test). These results demonstrate that endogenous neural precursors can be induced *in situ* to differentiate into mature neurons which survive for at least several months in the adult murine neocortex.

To examine the possibility that these new cells were expressing multiple phenotypic markers inappropriately, we triple-labelled sections with antibodies to GFAP, an astroglial marker, and MBP, an oligodendroglial marker. BrdU+/NeuN+ neurons were negative for GFAP (Fig. 1j) and MBP, confirming that they had differentiated specifically into mature neurons.

To understand further the anatomic source of the new neurons and whether proliferation of the precursors was altered, we examined the location of newly mitotic BrdU+ cells in experimental and control animals. After two weeks of BrdU treatment, BrdU+ non-neuronal cells were found throughout both experimental and control cortex. At two weeks, experimental mice ($n = 6$) had $5,400 \pm 1,500$ BrdU+ cells per mm^3 in the experimental regions, not significantly more than control mice ($n = 3$), which had $4,100 \pm 1,700$ BrdU+ cells per mm^3 . Such proliferation normally exists in the adult cortex; however, under normal circumstances,

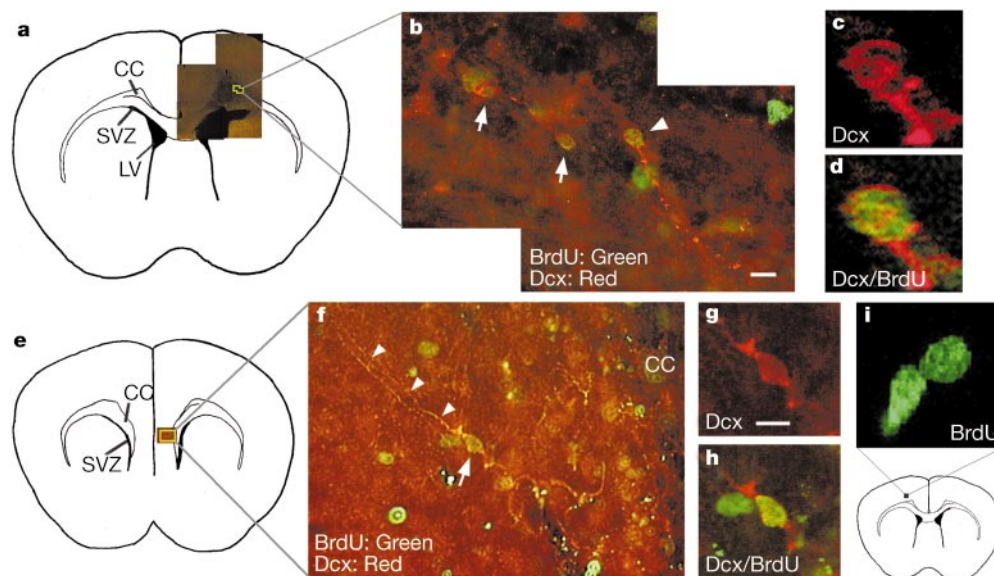


Figure 2 New cells express the migratory neuronal marker Doublecortin (Dcx) two weeks after induction of apoptosis. BrdU+/Dcx+ neurons appear to migrate from the SVZ, through the corpus callosum and into experimental regions of cortex. **a**, Camera lucida showing BrdU+ (green)/Dcx+ (red) neurons within the superficial CC. Lateral ventricle (LV). **b**, Fluorescence micrograph of three BrdU+/Dcx+ neurons. Scale bar: 25 μ m. **c,d**, Confocal micrographs confirm that the neuron in **b** (arrowhead) is BrdU+/Dcx+. **e**, Camera lucida: A BrdU+/Dcx+ neuron within cortex. **f**, Higher magnification: the

differentiating BrdU+/Dcx+ neuron (arrow) extending processes (arrowheads) in cortex. No BrdU+/Dcx+ cells were found in the CC or cortex of control mice. Dcx staining is specific for migrating neurons^{9,10}; it does not overlap with A2B5, O4, RIP, MBP or GFAP staining (see methods). **g, h**, Confocal micrographs of this new neuron confirm that it is BrdU+/Dcx+. Scale bar: 10 μ m. **i**, Pairs of BrdU+ non-neuronal cells were found throughout cortex in experimental and control mice.

mitotic cells differentiate into glia, remain undifferentiated, or die. Some new cells located within the cortex of experimental animals appear to originate from precursors located within cortex itself, whereas a second population, absent in control animals, appeared to originate in or near the subventricular zone. At two weeks, pairs of BrdU+ cells, apparently daughters of the same precursor, were found throughout both control and experimental cortex (Fig. 2i). Some of these new cells are possibly the daughters of the cortically located adult multipotent neural precursors that have been observed *in vitro*²². The adult subventricular zone (SVZ) and ependymal zone, located subjacent to the deep layers of cortex containing degenerating projection neurons, also contain a population of constitutively dividing, multipotent neural precursors^{2-4,23-26}, and appear to be the source of the second population of newborn cells. Cells from one or both of these populations may contribute to the neurogenesis reported here.

We investigated the early differentiation and migration of the new neurons using a marker of early postmitotic neurons, Doublecortin (Dcx). In experimental mice only, BrdU+/Dcx+ neurons with migratory morphologies appeared to migrate from the SVZ through the corpus callosum (Fig. 2a–d) and into experimental regions of cortex (Fig. 2e, f). Doublecortin is a microtubule-associated protein that is exclusively expressed in migrating and differentiating neurons, but is not detectable in mature neurons^{9,10}. Confocal analysis confirmed that these BrdU+ cells express Doublecortin (Fig. 2c, d). No BrdU+/Dcx+ cells were found in the corpus callosum or cortex of control animals, although they exist in the rostral migratory stream of both controls and experimental animals, confirming previous reports^{9,10}. BrdU+/Dcx+ neurons within the corpus callosum displayed morphologies characteristic of migrating neurons. New Dcx+ neurons located in deep layers of cortex displayed more complex morphologies with apical processes that indicate further differentiation. These results demonstrate the progressive differentiation of endogenous precursors: first into Dcx+ migratory neuroblasts, then into immature Dcx+ neurons with process extension, and then into more mature neurons.

We examined the continued differentiation of the new neurons using antibodies to Hu, an RNA-binding protein that begins to be expressed in neuronal nuclei and somata soon after differentiation^{11,12}. BrdU+/Hu+ neurons with large, ovoid nuclei were found in the cortex of experimental mice two through 28 weeks after induction of apoptosis (Fig. 3), and were absent in control mice. We used Hoechst 33258, which stains DNA, to confirm that a specific nucleus belonged to a particular cell. The dispersed heterochromatin pattern demonstrated by Hoechst staining coincident with the BrdU+ nuclei is characteristic of neuronal nuclei and confirms BrdU+/Hu+ double labelling. The expression of Hu confirms induced cortical neuron differentiation by endogenous neural precursors.

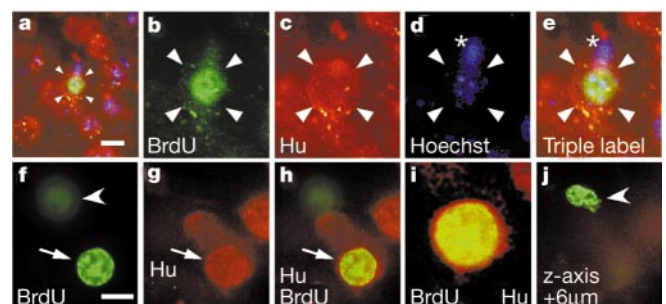


Figure 3 Newly generated BrdU+ cells express the early neuronal marker Hu in experimental cortex. **a**, A BrdU+/Hu+ neuron (arrowheads) at two weeks. Surrounding neurons are BrdU-negative. Scale bar: 10 μ m. **b–e**, Higher magnification. **b**, BrdU (green) labels the nucleus of a new neuron. **c**, Hu (red) labels the BrdU+ cell. **d**, Hoechst (blue) shows the large, dispersed nucleus of the new neuron and the compact nucleus of an adjacent glial cell (asterisk). **e**, Overlay. **f**, Arrow indicates BrdU+ (green) nucleus of a BrdU+/Hu+ neuron at 28 weeks. Arrowhead indicates a BrdU+/Hu- cell in another plane of focus. **g**, Hu+ soma (red). **h**, Overlay. **i**, Confocal confirmation of double labelling. **j**, A BrdU+/Hu- cell located in a focal plane 6 μ m higher demonstrates specificity of staining.

To define the identity of the BrdU+ neurons further and to examine their ability to establish long-distance connections, we injected the retrograde label FluoroGold into the same thalamic sites as the original nanosphere injections. Previous experiments demonstrate that large numbers of embryonic neurons transplanted into regions of adult cortex undergoing apoptotic degeneration of projection neurons can re-form long distance connections to the original targets of the degenerating neurons¹⁹. Here, at nine weeks we observed BrdU+, FluoroGold-labelled neurons that had large nuclei and large cell bodies indicating pyramidal projection neuron morphology (Fig. 4). We estimate that several per cent of new neurons send long-distance projections to the thalamus. These results show that endogenous precursors that differentiate into mature neurons can establish appropriate, long-distance corticothalamic connections in the adult brain.

These results demonstrate that endogenous neural precursors can be induced to differentiate into neocortical neurons in a layer- and region-specific manner, and can re-form appropriate corticothalamic connections, in regions of adult mammalian neocortex that do not normally undergo neurogenesis. Although we observed a significant number of new neurons, the BrdU labelling probably underestimates this number because BrdU is only available for nuclear incorporation for a few hours after treatment. It is possible that the small number of available endogenous neural precursors also limits the amount of neurogenesis. In future experiments, it would be of interest to examine the effects of expanding the number of multipotent precursors in the SVZ by treatment with mitogenic growth factors^{27,28}.

In these experiments, we induced the death of a much lower proportion of targeted deep layer VI neurons (approximately 20%)

compared with our previous experiments, in which we induced the death of 60%–90% of more superficial targeted neurons in neocortical layer II/III in mice, or in area HVC of the avian forebrain. It is unlikely that the new neurons we observed incorporated BrdU whilst undergoing apoptosis, because many labelled neurons were present in neocortex months after the end of the BrdU treatment and we were able to track the developmental progression of these cells using Doublecortin.

Our results support the view that adult cortex is capable of neuronal regeneration in response to neuronal injury or degeneration, under specific and appropriate conditions. Precedents exist by which neurogenesis can be increased in regions where it is already occurring^{21,27–30}. In these instances, it might be sufficient to increase proliferation or survival of precursors, so additional competent precursors can respond to differentiation signals that already exist. In contrast, no precedents exist for inducing neurogenesis of projection neurons *de novo* in the neocortex, where it does not normally occur. For induction of *de novo* neocortical neurogenesis, we propose that restrictions may be removed and/or that a program of developmental gene expression may be reactivated. The synchronous targeted apoptosis induced here may be uncovering a normal cellular and molecular response to neuronal death that has not been previously observed because of its low level of activation with sporadic neuronal death. Alternatively, targeted apoptosis may be releasing a repressed program of neural plasticity that permits and/or instructs early neocortical differentiation. Our results indicate that it may be possible to manipulate endogenous neural precursors *in situ* to undergo neurogenesis in the adult brain. Elucidation of the relevant molecular controls may allow the development of neuronal replacement therapies for neurodegenerative disease and other CNS injury that do not require transplantation of exogenous cells. □

Methods

Induction of apoptosis

We conjugated chlorin *e*₆ to subfractionated nanospheres (Lumafluor), as described⁷. We stereotactically injected these bilaterally at three sites directed at nuclei VLc, VPLo, VLo and Area X of thalamus in adult female mice (> 6 weeks), from which they were retrogradely transported to the somata of layer VI corticothalamic projection neurons. We made the injections via a posterior approach that did not come near anterior cortex¹⁹, so there was no nonspecific surgical injury to the region of cortex where we later induced targeted neuronal degeneration. Two weeks later, we exposed the anterior cortex through intact dura to 674 nm wavelength light from a solid-state, continuous wave laser using custom optics to collimate the light at the level of layer VI. Photoactivated chlorin *e*₆ produced singlet oxygen within neuronal lysosomes, inducing apoptosis exclusively in nanosphere-containing projection neurons^{7,8,18}. Degeneration of about 20% of targeted projection neurons occurred, with neuronal degeneration confirmed in previous studies^{7,8,21} by loss of retrogradely labelled neurons compared with adjacent control regions, by the presence of TUNEL+ nuclei in projection neurons, confirming apoptotic death, and by analysis of routine histologic sections one month later.

BrdU labelling of dividing cells

We gave experimental and control mice either twice-daily BrdU injections (120 mg kg⁻¹) or drinking water containing BrdU (1.5 mg ml⁻¹) for the two weeks following induction of apoptosis. Control mice had no BrdU+ neurons in cortex, demonstrating that BrdU was not incorporated into mature neurons owing to toxicity.

Immunocytochemistry

To exclude the possibility of false labelling, we used primary antibodies produced in different species and secondary antibodies tested for cross reactivity in all experiments. We washed paraformaldehyde-fixed sections (40 µm) in HCl for 2 h at room temperature; blocked them in PBS containing 4% goat serum, 0.3% BSA, and 0.3% triton; and then incubated them with primary antibodies overnight and secondary antibodies for 2 h in blocking solution. We incubated the sections in rat anti-BrdU (1:100, Harlan) and the following primary antibodies: rabbit anti-Doublecortin (1:100, courtesy of J. G. Gleeson and C. A. Walsh); mouse anti-Hu (1:50, courtesy of S. A. Goldman and M. Marusich); mouse anti-NeuN antibody (1:100, Chemicon); rabbit anti-GFAP (1:100, Sigma); and mouse anti-MBP (1:1000, Sternberger). We used highly cross-adsorbed anti-mouse, anti-rat and anti-rabbit Alexa (Molecular Probes) secondary antibodies to avoid cross reactivity. We used Hoechst 33258 as a nuclear stain.

We double-labelled early postnatal mouse brain with antibodies to Dcx and the glial or glial precursor markers A2B5 (1:500, courtesy of P. Follett), O4 (1:1000, courtesy of P. Follett), RIP (1:200,000, Chemicon), MBP or GFAP. No Dcx+ cells expressed any of

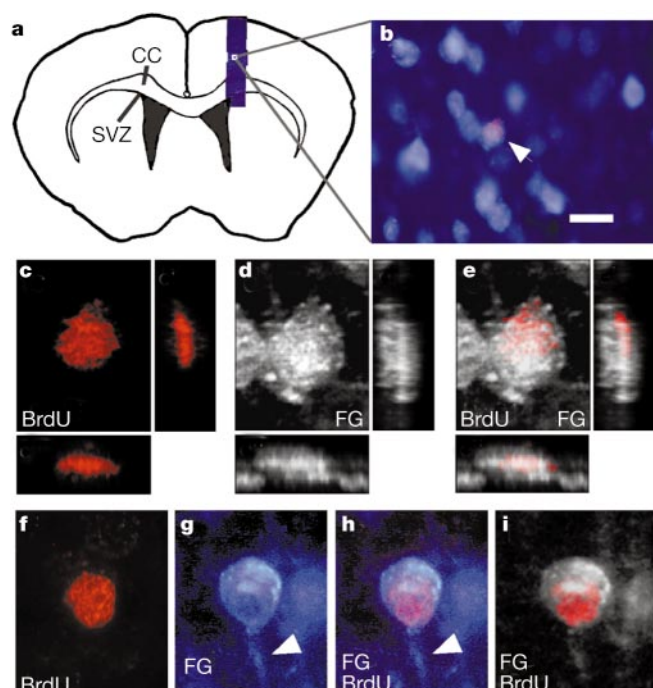


Figure 4 Newly generated BrdU+ neurons in experimental cortex extend long-distance projections to thalamus. **a**, Camera lucida of a BrdU+ (red) / FG+ (white) retrogradely labelled neuron in layer VI of cortex at nine weeks. **b**, Higher magnification shows the new corticothalamic neuron (arrow) and FG+/BrdU– original neurons. Scale bar: 20 µm. **c–e**, Confocal 3D reconstruction of the neuron. Right, x-axis; bottom, y-axis. **c**, BrdU+ nucleus, indicating the cell underwent mitosis in the two weeks following induction of apoptosis. **d**, FluoroGold+ soma, indicating that this neuron projects to thalamus. **e**, Merged image, confirming double-labelling in three dimensions. **f**, BrdU+ (red) nucleus of a new FG+ corticothalamic neuron. **g**, FG+ (blue) cell body with labelled axon (arrow). **h**, Overlay. **i**, Confocal microscopy confirms double labelling.

these glial markers during this peak period of gliogenesis, confirming that Dcx is expressed solely in immature neurons.

Microscopy

We performed confocal microscopy using a Noran Instruments Oz Confocal Laser Scanning Microscope, a Nikon Diaphot optical microscope, and Intervision 3D analysis software. We produced 3D digital reconstructions from a series of confocal images taken at 0.3 or 0.5 μm intervals through the region of interest. This 3D analysis allowed us unequivocally to co-localize BrdU and markers of differentiation. We compiled the montage in Fig. 1f from three sequential sections from the anterior forebrain of an experimental 28-week-survival mouse.

Quantification

We quantified the number of BrdU+ cells and BrdU+/NeuN+ neurons in experimental and control regions by sampling every sixth section, using a modified version of the fractionator method²¹. The experimental regions of cortex extended from dorsal cortex to the medial bank, and spanned the thickness of cortical layer VI. We compared numbers of new neurons in experimental mice to control mice by multiple statistical methods: unpaired two-tailed *t* test; non-parametric two-tailed Mann–Whitney test; and unpaired *t*-test with Welch correction. All methods yielded *P* values less than 0.0009, demonstrating a highly significant difference between the groups.

FluoroGold Injections

We stereotactically injected 32 nl of 3% FluoroGold solution dissolved in water bilaterally at each of the same three thalamic sites as the original nanosphere injections. We injected the retrograde label FluoroGold into the thalamus of adult mice eight weeks after induction of targeted corticothalamic neuron apoptosis, and perfused the mice at nine weeks.

Received 11 January; 3 May 2000.

- Altman, J. & Das, G. D. Autoradiographic and histological evidence of postnatal hippocampal neurogenesis in rats. *J. Comp. Neurol.* **124**, 319–335 (1965).
- Altman, J. Autoradiographic and histological studies of postnatal neurogenesis. IV. Cell proliferation and migration in the anterior forebrain, with special reference to persisting neurogenesis in the olfactory bulb. *J. Comp. Neurol.* **137**, 433–457 (1969).
- Luskin, M. B. Restricted proliferation and migration of postnatally generated neurons derived from the forebrain subventricular zone. *Neuron* **11**, 173–189 (1993).
- Lois, C. & Alvarez-Buylla, A. Proliferating subventricular zone cells in the adult mammalian forebrain can differentiate into neurons and glia. *Proc. Natl Acad. Sci. USA* **90**, 2074–2077 (1993).
- Kaplan, M. S. & Hinds, J. W. Neurogenesis in the adult rat: electron microscopic analysis of light radioautographs. *Science* **197**, 1092–1094 (1977).
- Gould, E., Reeves, A. J., Graziano, M. S. A. & Gross, C. G. Neurogenesis in the neocortex of adult primates. *Science* **286**, 548–552 (1999).
- Macklis, J. D. Transplanted neocortical neurons migrate selectively into regions of neuronal degeneration produced by chromophore-targeted laser photolysis. *J. Neurosci.* **13**, 3848–3863 (1993).
- Sheen, V. L. & Macklis, J. D. Targeted neocortical cell death in adult mice guides migration and differentiation of transplanted embryonic neurons. *J. Neurosci.* **15**, 8378–8392 (1995).
- Gleeson, J. G., Lin, P. T., Flanagan, L. A. & Walsh, C. A. Doublecortin is a microtubule-associated protein and is expressed widely by migrating neurons. *Neuron* **23**, 257–271 (1999).
- Francis, F. et al. Doublecortin is a developmentally regulated, microtubule-associated protein expressed in migrating and differentiating neurons. *Neuron* **23**, 247–256 (1999).
- Marusch, M. F., Furneaux, H. M., Henion, P. D. & Weston, J. A. Hu neuronal proteins are expressed in proliferating neurogenic cells. *J. Neurobiol.* **25**, 143–155 (1994).
- Barami, K., Iversen, K., Furneaux, H. & Goldman, S. A. Hu protein as an early marker of neuronal phenotypic differentiation by subependymal zone cells of the adult songbird forebrain. *J. Neurobiol.* **28**, 82–101 (1995).
- Reh, T. A. Cell-specific regulation of neuronal production in the larval frog retina. *J. Neurosci.* **7**, 3317–3324 (1987).
- Parent, J. M. et al. Dentate granule cell neurogenesis is increased by seizures and contributes to aberrant network reorganization in the adult rat hippocampus. *J. Neurosci.* **17**, 3727–3738 (1997).
- Snyder, E. Y., Yoon, C., Flax, J. D. & Macklis, J. D. Multipotent neural precursors can differentiate toward replacement of neurons undergoing targeted apoptotic degeneration in adult mouse neocortex. *Proc. Natl Acad. Sci. USA* **94**, 11663–11668 (1997).
- Kirn, J. R. & Nottebohm, F. Direct evidence for loss and replacement of projection neurons in adult canary brain. *J. Neurosci.* **13**, 1654–1663 (1993).
- Gould, E. & Cameron, H. A. Regulation of neuronal birth, migration and death in the rat dentate gyrus. *Dev. Neurosci.* **18**, 22–35 (1996).
- Wang, Y., Sheen, V. L. & Macklis, J. D. Cortical interneurons upregulate neurotrophins *in vivo* in response to targeted apoptotic degeneration of neighboring pyramidal neurons. *Exp. Neurol.* **154**, 389–402 (1998).
- Hernit-Grant, C. S. & Macklis, J. D. Embryonic neurons transplanted to regions of targeted photolytic cell death in adult mouse somatosensory cortex re-form specific callosal projections. *Exp. Neurol.* **139**, 131–142 (1996).
- Shin, J. S. Transplanted neuroblasts differentiate appropriately into projection neurons with the correct neurotransmitter and receptor phenotype in neocortex undergoing targeted projection neuron neurodegeneration. *J. Neurosci.* (in the press).
- Scharff, C. et al. Targeted neuronal death affects neuronal replacement and vocal behavior in adult songbirds. *Neuron* **25**, 481–492 (2000).
- Palmer, T. D. et al. Fibroblast growth factor-2 activates a latent neurogenic program in neural stem cells from diverse regions of the adult CNS. *J. Neurosci.* **19**, 8487–8497 (1999).
- Reynolds, B. A. & Weiss, S. Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. *Science* **255**, 1707–1710 (1992).

- Richards, L. J., Kilpatrick, T. J. & Bartlett, P. F. *De novo* generation of neuronal cells from the adult mouse brain. *Proc. Natl Acad. Sci. USA* **89**, 8591–8595 (1992).
- Vescovi, A. L., Reynolds, B. A., Fraser, D. D. & Weiss, S. bFGF regulates the proliferative fate of unipotent (neuronal) and bipotent (neuronal/astroglial) EGF-generated CNS progenitor cells. *Neuron* **11**, 951–966 (1993).
- Morshead, C. M. et al. Neural stem cells in the adult mammalian forebrain: a relatively quiescent subpopulation of subependymal cells. *Neuron* **13**, 1071–1082 (1994).
- Craig, C. G. et al. *In vivo* growth factor expansion of endogenous subependymal neural precursor cell populations in the adult mouse brain. *J. Neurosci.* **16**, 2649–2658 (1996).
- Kuhn, H. G. et al. Epidermal growth factor and fibroblast growth factor-2 have different effects on neural progenitors in the adult rat brain. *J. Neurosci.* **17**, 5820–5829 (1997).
- Zigova, T., Pencea, V., Wiegand, S. J. & Luskin, M. B. Intraventricular administration of BDNF increases the number of newly generated neurons in the adult olfactory bulb. *Mol. Cell. Neurosci.* **11**, 234–245 (1998).
- Cameron, H. A., Hazel, T. G. & McKay, R. D. Regulation of neurogenesis by growth factors and neurotransmitters. *J. Neurobiol.* **36**, 287–306 (1998).
- Guillery, R. W. & Herrup, K. Quantification without pontification: choosing a method for counting objects in sectioned tissues. *J. Comp. Neurol.* **386**, 2–7 (1997).

Acknowledgements

We thank J. Gleeson, S. Goldman, M. Marusch, C. Walsh, and P. Follett for reagents; M. Christian and C. Tai for technical assistance; L. Catapano, P. Follett, R. Fricker, M. Gates, and J. Gleeson for discussions; and L. Benowitz, M. Greenberg, L. Kunkel, and C. Walsh for critical reading. This work was supported by grants from the NIH (J.D.M.), Alzheimer's Association (J.D.M.), Human Frontiers Science Program (J.D.M.), an NIH predoctoral training grant (S.S.M.), and a postdoctoral fellowship from the Canadian MRC (B.R.L.).

Correspondence and requests for materials should be addressed to J.D.M. (e-mail: macklis@hub.tch.harvard.edu).

Regulation of distinct AMPA receptor phosphorylation sites during bidirectional synaptic plasticity

Hey-Kyoung Lee*, Michaela Barbarosie†, Kimihiko Kameyama‡, Mark F. Bear† & Richard L. Huganir*

* Howard Hughes Medical Institute, Department of Neuroscience, Johns Hopkins Medical School, Baltimore, Maryland 1205, USA

† Howard Hughes Medical Institute, Department of Neuroscience, Brown University, Providence, Richmond 02912, USA

‡ Laboratory of Molecular Neurobiology, National Institute of Bioscience and Human Technology, Ibaraki 305-8566, Japan

Bidirectional changes in the efficacy of neuronal synaptic transmission, such as hippocampal long-term potentiation (LTP) and long-term depression (LTD), are thought to be mechanisms for information storage in the brain^{1–4}. LTP and LTD may be mediated by the modulation of AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid) receptor phosphorylation^{5–7}. Here we show that LTP and LTD reversibly modify the phosphorylation of the AMPA receptor GluR1 subunit. However, contrary to the hypothesis that LTP and LTD are the functional inverse of each other, we find that they are associated with phosphorylation and dephosphorylation, respectively, of distinct GluR1 phosphorylation sites. Moreover, the site modulated depends on the stimulation history of the synapse. LTD induction in naive synapses dephosphorylates the major cyclic-AMP-dependent protein kinase (PKA) site, whereas in potentiated synapses the major calcium/calmodulin-dependent protein kinase II (CaMKII) site is dephosphorylated. Conversely, LTP induction in naive synapses and depressed synapses increases phosphorylation of the CaMKII site and the PKA site, respectively. LTP is differentially sensitive to CaMKII and PKA inhibitors depending on the history of the synapse. These results indicate that AMPA receptor phosphorylation is critical for synaptic plasticity, and that identical stimula-

tion conditions recruit different signal-transduction pathways depending on synaptic history.

AMPA receptors, the principal excitatory neurotransmitter receptors in the brain, are composed of four different subunits (GluR1–4) that can assemble in different combinations^{8,9}. The GluR1 subunit is predominantly expressed in the forebrain, including the hippocampus, a region implicated in memory formation¹⁰. A study using GluR1 knockout mice has shown that this subunit is essential for LTP in the CA1 region of the adult hippocampus¹¹. The GluR1 subunit is regulated by protein phosphorylation at two sites on its intracellular carboxy-terminal domain¹². Serine 831 is phosphorylated by CaMKII and protein kinase C (PKC), whereas serine 845 is phosphorylated by PKA^{12–14}. Phosphorylation of either of these sites potentiates AMPA receptor function through distinct biophysical mechanisms^{12,13,15,16}.

The aim of this study was to examine changes in phosphorylation of GluR1 at these two sites after the synaptic induction of LTP and LTD using phosphorylation-site-specific antibodies. A chemically induced form of LTD (chemLTD) is associated with persistent dephosphorylation of GluR1 at a PKA phosphorylation site, Ser 845 (refs 6, 7). The chemLTD approach was designed to maximize the number of affected synapses, thereby increasing the probability of detecting biochemical changes. We have been able to increase the sensitivity of our assay to detect small changes in GluR1 phosphorylation on Ser 831 and Ser 845 in single hippocampal slices after the synaptic induction of LTP and LTD.

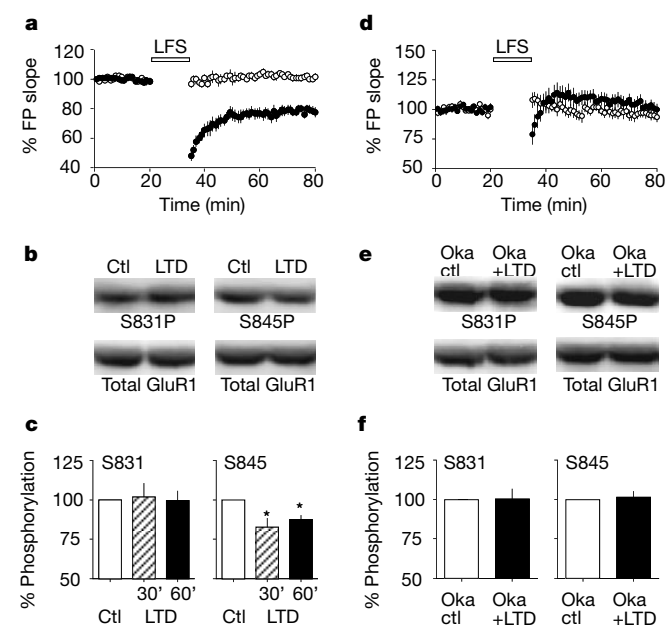


Figure 1 Homosynaptic LTD in CA1 is associated with dephosphorylation of GluR1 at a PKA site. **a**, Simultaneous recording of slices receiving baseline stimulation (control, open circles) and 1-Hz stimulation (LTD, closed circles; $n = 17$ each). FP, field potential. **b**, Immunoblot using antibody against phosphorylated Ser 831 (CaMKII site; S831P) and phosphorylated Ser 845 (PKA site; S845P) on GluR1. Both blots were stripped and reprobed with the antibody recognizing the C-terminal end of GluR1 (total GluR1). **c**, Summary of phosphorylation changes in Ser 831 and Ser 845 following LTD induction 30 min ($n = 11$) and 1 h ($n = 17$) after LFS. Asterisks indicate value statistically significantly different from control. **d**, Homosynaptic LTD is blocked by pretreatment of slices with okadaic acid (control slices: Oka ctl, open circles; LTD slices: Oka+LTD, filled circles; $n = 11$ each). **e**, Immunoblot of slices pretreated with okadaic acid using phosphorylation-site-specific antibodies. Panels as in **b**. Okadaic acid pretreatment increased the phosphorylation of Ser 831 by $155 \pm 17\%$ and Ser 845 by $139 \pm 24\%$. **f**, Okadaic acid pretreatment prevented the dephosphorylation of Ser 845 (PKA site) associated with homosynaptic LTD. The signals were normalized as described in **c**.

To analyse phosphorylation of GluR1 during synaptic plasticity, we placed two hippocampal slices in the same recording chamber and recorded extracellular field potentials simultaneously in the CA1 dendritic region of both slices upon stimulation of the Schaffer collaterals. After collecting a stable baseline, stimulation to one slice was turned off while the other slice received low-frequency stimulation (LFS; 1 Hz, 900 pulses). After the LFS, the stimulation was returned to the baseline frequency, and stimulation to the control slice was resumed. The slice that received LFS showed homosynaptic LTD ($78 \pm 3\%$ of baseline field potential slope 1 h after the onset of 1-Hz stimulation, $n = 17$), but there was no significant change in synaptic strength in the control slice ($101 \pm 2\%$ of baseline, $n = 17$; Fig. 1a). One hour after the induction of LTD the slices were collected and frozen immediately on dry ice. We then analysed the phosphorylation states of GluR1 at Ser 831 and Ser 845 in the control and LTD slices by quantitative immunoblotting.

Like chemLTD, homosynaptic LTD produced specific dephosphorylation of Ser 845 ($87 \pm 3\%$ phosphorylation of control slices, $n = 17$; paired t -test: $P < 0.01$; Fig. 1c), whereas there was no significant change in phosphorylation of Ser 831 ($99 \pm 6\%$ of control; Fig. 1b and c). The dephosphorylation at Ser 845 could be detected as early as 30 min after the onset of LFS ($83 \pm 5\%$ of control, $n = 11$; Fig. 1c). As expected, the magnitude of the change in GluR1 phosphorylation following homosynaptic LTD was smaller than that reported for chemLTD, because only a small percentage of synapses in the slice are depressed during LFS-induced homosynaptic LTD. This effect was, however, reproducible and statistically significant. In addition, there was no dephosphorylation in slices that did not exhibit LTD after LFS, either because induction of LTD failed or because the NMDA (*N*-methyl-D-aspartate) receptor antagonist AP5 (*D*(-)-2-amino-5-phosphonovaleric acid) was

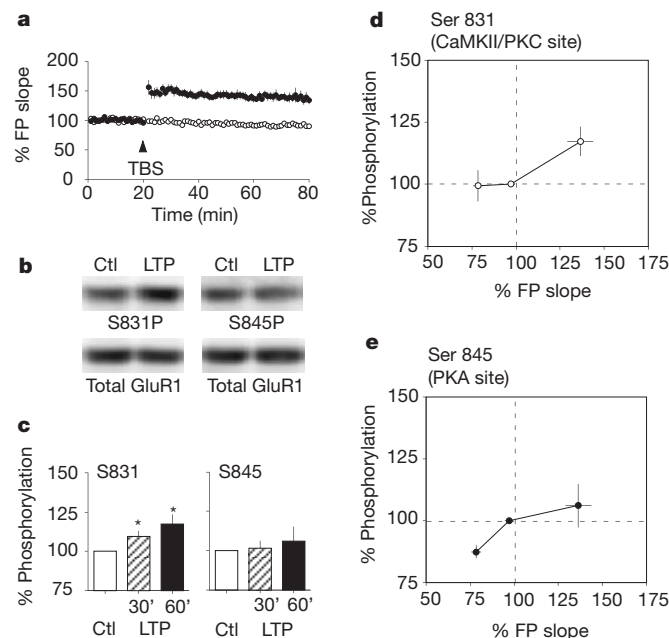


Figure 2 LTP induction increases phosphorylation at CaMKII phosphorylation site on GluR1. **a**, Simultaneous recording of field potentials in control slices (open circles) and LTP slices (closed circles; $n = 13$ each). **b**, Immunoblots using phosphorylation-site-specific antibodies. Panels as in Fig. 1b. **c**, Quantification of immunoblots with samples from control slices (Ctl) and 30 min ($n = 9$) and 1 h ($n = 13$) after LTP induction. One data point was excluded from analysis, as the phosphorylation changes in Ser 845 exceeded 2 s.d. from the mean. The inclusion of this data point does not change the conclusions. **d**, Changes in phosphorylation of GluR1 at the CaMKII site (S831) 1 h after induction of LTP and LTD. **e**, GluR1 phosphorylation changes at the PKA site (S845) associated with bidirectional synaptic plasticity 1 h after induction.

applied (data not shown). Collectively, these results confirm that chemLTD and homosynaptic LTD share similar downstream expression mechanism(s)⁶.

Homosynaptic LTD is dependent on postsynaptic protein phosphatase activity^{17,18}. To test whether the dephosphorylation of GluR1 following LTD is also sensitive to protein phosphatase inhibitors, we incubated the control and experimental slices in okadaic acid (1 μ M in ACSF with 0.1% dimethyl sulphoxide (DMSO)) for 2–3 h and transferred them to the recording chamber. As reported, okadaic acid abolished LTD ($101 \pm 4\%$ of baseline, $n = 11$, Fig. 1d; LTD in vehicle solution: $76 \pm 5\%$, $n = 10$, data combined in Fig. 1a; t -test, $P < 0.01$). Moreover, pretreatment of slices with okadaic acid blocked the LFS-induced dephosphorylation of Ser 845 ($101 \pm 7\%$ of OKA controls; $n = 11$; Fig. 1e and f). In contrast, induction of LTD in slices kept in vehicle solution still produced significant dephosphorylation at Ser 845 ($88 \pm 3\%$ of controls; $n = 10$; paired t -test, $P < 0.05$; data combined in Fig. 1c). This result indicates that protein phosphatase 1/2A may be critical for the LFS-induced dephosphorylation of GluR1 at Ser 845.

LTP is probably associated with an increase in phosphorylation of GluR1 by CaMKII (ref. 5). We tested whether we could detect the same changes using our phosphorylation-site-specific antibodies. Slices that received theta burst stimulation (TBS) showed LTP ($134 \pm 3\%$ of baseline at 1 h after TBS, $n = 14$; Fig. 2a), and a significant increase in GluR1 phosphorylation at Ser 831 (CaMKII/PKC site) at both 30 min ($110 \pm 3\%$ of controls, $n = 9$; paired t -test, $P < 0.05$) and 1 h ($117 \pm 5\%$ of controls, $n = 14$; paired t -test, $P < 0.05$; Fig. 2c) after TBS. Phosphorylation of Ser 845 (PKA site) was not significantly increased 30 or 60 min after LTP induction (Fig. 2b and c). In slices that did not exhibit LTP, GluR1 phosphorylation did not change (data not shown). These results agree

with those of previous studies indicating that the early phase of LTP is dependent on postsynaptic CaMKII activation^{19,22} and that phosphorylation of GluR1 by CaMKII increases following LTP⁵.

The data presented so far indicate that LTP and LTD may be associated with changes in phosphorylation of GluR1 at different sites (Fig. 2d and e). LTD is associated with dephosphorylation at Ser 845, a PKA site, whereas LTP is associated with increased phosphorylation of Ser 831, a CaMKII site. This indicates that although induction of LTP and LTD results in the bidirectional control of AMPA-receptor phosphorylation, LTP and LTD do not regulate phosphorylation of the same site. As LTP and LTD are reversible processes, we tested what happens to GluR1 phosphorylation after the reversal of LTP with LFS ('depotentiation') and the reversal of LTD with TBS ('de-depression').

To examine depotentiation we subjected both slices in the recording chamber to TBS simultaneously, which resulted in LTP. After 30 min of recording, LFS was given to one of the slices. This stimulation depressed the synaptic strength back to the baseline level (depotentiation; $93 \pm 4\%$ of baseline 30 min after onset of 1-Hz stimulation, $n = 14$; Fig. 3a) while the synaptic response stayed potentiated in slices that only received TBS ($139 \pm 3\%$ of baseline at 1 h post-TBS; Fig. 3a). There was significant dephosphorylation of Ser 831 in depotentiated slices ($89 \pm 5\%$ of LTP slices, $n = 14$; paired t -test: $P < 0.05$) but no significant change in Ser 845 ($106 \pm 5\%$ of LTP slices; Fig. 3b and c). This result indicates that the same stimulation protocol (LFS) results in the dephosphorylation of different sites on GluR1 depending on the previous experience of the synapse (whether it had previously undergone LTP).

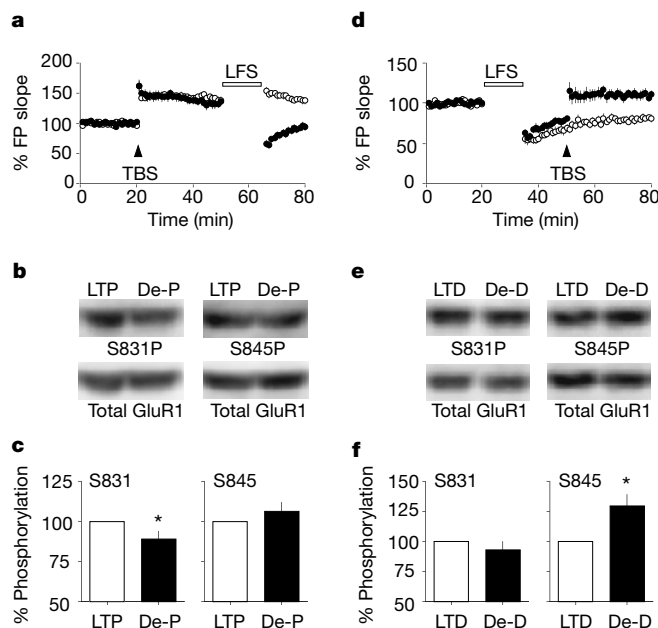


Figure 3 Depotentiation results in dephosphorylation of a CaMKII site on GluR1, whereas de-depression is associated with an increase in phosphorylation of a PKA site. **a**, To detect changes associated with depotentiation (De-P, closed circles), we compared it with LTP (open circles; $n = 14$ each). **b**, Representative immunoblots of LTP and depotentiation (De-P) samples. Panels as in Fig. 1b. **c**, Depotentiation dephosphorylates Ser 831 (CaMKII site). Open bars, phosphorylation levels in LTP slices; closed bars, per cent phosphorylation in depotentiation slices (De-P; $n = 14$ each). **d**, De-depression (De-D, closed circles) was tested against LTD (open circles; $n = 10$ each). **e**, Immunoblots of LTD and de-depression (De-D) slices. **f**, De-depression increases phosphorylation at Ser 845 (PKA site).

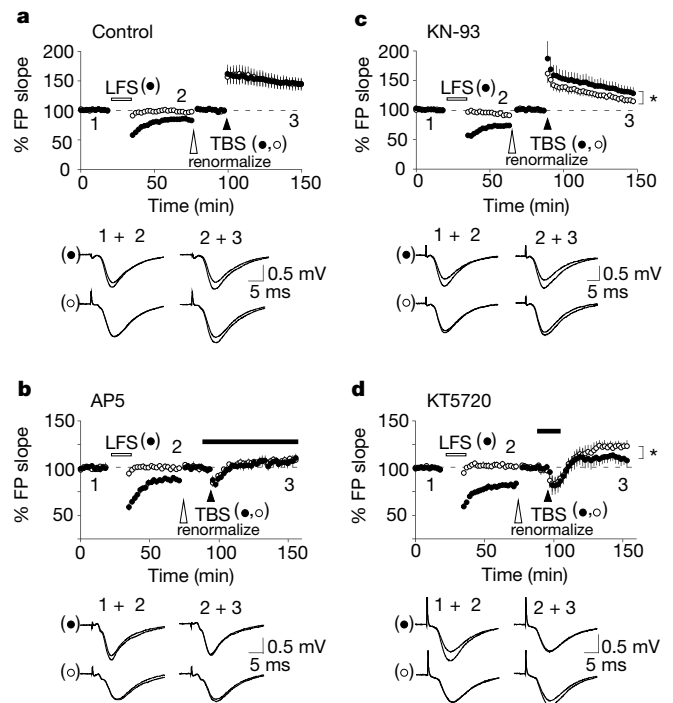


Figure 4 De-depression and LTP utilize different signal transduction pathways for their expression. **a**, De-depression (closed circles) and LTP (open circles) in normal ACSF ($n = 11$) in a two-pathway experiment. Sample field potential responses from one experiment taken at time points indicated are shown below. **b**, Both de-depression (closed symbols) and LTP (open symbols) are blocked by DL-AP5 (100 μ M; $n = 4$). Below, field potential recordings from a typical experiment. **c**, The CaMKII inhibitor KN93 affects LTP (open circles) more than de-depression (closed circles; $n = 8$). Below, examples of field potential responses from one experiment. **d**, PKA inhibitor KT5720 reduced de-depression (closed circles) more than LTP (open circles; $n = 6$). KT5720 (1 μ M) was bath applied for 10 min during which TBS was delivered to both pathways. Below, representative field potential recordings.

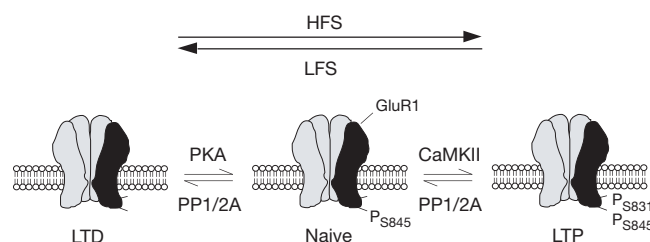


Figure 5 A model explaining the bidirectional changes in AMPA-receptor phosphorylation and NMDA-receptor-dependent synaptic plasticity. High-frequency stimulation (HFS) delivered to naive synapses preferentially activates CaMKII, which increases phosphorylation of GluR1 on Ser 831 (CaMKII site; P₈₃₁) resulting in LTP. In contrast, low frequency stimulation (LFS) given in a naive slice activates protein phosphatases

(including PP1/2A), which dephosphorylate Ser 845 (PKA site; P₈₄₅). However, LFS given to a previously potentiated synapse (LTP) dephosphorylates Ser 831, as shown in our depotentiation result. On the other hand, HFS delivered to previously depressed synapses (LTD) phosphorylates Ser 845, as shown by our de-depression data.

To examine de-depression, we induced LTD in two slices in the recording chamber; 30 min later we delivered TBS to one of the slices. The synaptic response in slices that received TBS was potentiated back to around the baseline response ($110 \pm 4\%$ of baseline, $n = 10$; Fig. 3d) whereas the synaptic strength in slices that only received LFS stayed depressed throughout the experiment ($81 \pm 4\%$ of baseline; Fig. 3d). There was no significant change in phosphorylation of Ser 831 (CaMKII/PKC site) following de-depression ($93 \pm 6\%$ of LTD slices, $n = 10$; Fig. 3e and f). Interestingly, the de-depression increased phosphorylation at Ser 845 (PKA site; $129 \pm 9\%$ of LTD slices; paired t -test: $P < 0.05$; Fig. 3e and f). These results indicate that, like LTD-inducing stimuli, LTP-inducing stimuli result in the differential regulation of GluR1 phosphorylation depending on the previous experience of the synapse. TBS delivered to 'naive' synapses causes phosphorylation of a CaMKII site, but TBS delivered to previously depressed synapses causes phosphorylation of a PKA site.

From these results we predict that the synaptic potentiation due to changes in AMPA receptor phosphorylation following TBS-induced de-depression or TBS-induced LTP should be differentially sensitive to CaMKII and PKA inhibitors. To test this hypothesis, we performed two-pathway experiments in which we could directly compare, in the same slices, the effects of TBS on naive and previously depressed synaptic inputs (Fig. 4). Under control conditions, TBS produced the same amount of synaptic potentiation regardless of the initial state of the synapses ($n = 11$; Fig. 4a), and both LTP and de-depression were prevented when NMDA receptors were blocked ($n = 4$; Fig. 4b). However, when slices were incubated in the selective CaMKII inhibitor KN-93 ($10 \mu\text{M}$ in 0.1% DMSO), TBS produced significantly less *de novo* potentiation than de-depression (measured 50 min after TBS; $P < 0.02$, paired t -test; $n = 8$; Fig. 4c). Conversely, when PKA was transiently inhibited at the time of tetanic stimulation using the selective inhibitor KT5720 ($1 \mu\text{M}$ in 0.1% DMSO), TBS produced significantly less de-depression than *de novo* LTP (measured 50 min after TBS; $P < 0.02$, paired t -test; $n = 6$; Fig. 4d). Thus, the relative contributions of CaMKII and PKA to TBS-induced potentiation vary depending on the initial state of the synapse, as predicted. This conclusion agrees with a previous report suggesting that there are different LTP-expression mechanisms in naive versus depressed synapses in PKC- γ knockout mice²³.

Our results show that bidirectional changes in synaptic function are associated with the reversible regulation of AMPA receptor phosphorylation. However, the phosphorylation or dephosphorylation of GluR1 occurred on distinct sites, depending on the past experience of the synapse (Fig. 5). Although the mechanism of this differential phosphorylation remains to be determined, it may involve differential activation of protein kinases and protein phosphatases in the potentiated, naive and depressed synapses. This differential bidirectional regulation of the phosphorylation of

AMPA receptors has significant computational implications for synaptic plasticity, as the results indicate that there may be at least three states of AMPA receptor activity with limited and regulated transitions between the states.

Although phosphorylation of the GluR1 subunit on either Ser 831 or Ser 845 potentiates AMPA receptor function, it appears to do so through distinct biophysical mechanisms^{15,16}. PKA phosphorylation of Ser 845 regulates the open channel probability of AMPA receptors¹⁶, whereas CaMKII phosphorylation regulates the apparent single-channel conductance of the receptor¹⁵. Note that single-channel conductance increases with LTP²⁴. Regulation of phosphorylation at these two sites should contribute to the changes in the efficacy of synaptic transmission that are observed during LTP and LTD. However, the trafficking and synaptic targeting of AMPA receptors may also be important in LTP and LTD^{25,26}. The role of phosphorylation of the GluR1 subunit in the regulation of AMPA receptor synaptic targeting is unclear. It is possible that phosphorylation of these sites regulates synaptic trafficking of the AMPA receptor in addition to regulating channel function. Alternatively, the synaptic trafficking of AMPA receptors may be regulated through a distinct mechanism, occurring with and complementing the modification of channel properties. Our findings that different phosphorylation sites and signal-transduction pathways contribute to bidirectional synaptic modifications in the hippocampus provide evidence for unexpected complexity in the control of the gain of excitatory synaptic transmission. □

Methods

Slice preparation

We prepared hippocampal slices from postnatal day 21–30 male Long–Evans rats (Harlan) as described⁶. In brief, dissection was done using ice-cold dissection buffer (composition in mM: sucrose, 212.7; KCl, 2.6; NaH₂PO₄, 1.23; NaHCO₃, 26; dextrose, 10; MgCl₂, 3; CaCl₂, 1) bubbled with 95% O₂/5% CO₂. A block of hippocampus was removed and sectioned into 400- μm -thick slices using a vibratome. Area CA3 was surgically removed after sectioning. The slices were transferred to a submersion-type holding chamber containing artificial cerebrospinal fluid (ACSF; in mM: NaCl, 124; KCl, 5; NaH₂PO₄, 1.25; NaHCO₃, 26; dextrose, 10; MgCl₂, 1.5; CaCl₂, 2.5) bubbled with 95% O₂/5% CO₂ and equilibrated at room temperature for about 1 h. Slices were then transferred to a submersion-type recording chamber continually perfused with 29–30 °C oxygenated ACSF at 2 ml min⁻¹. For the experiments shown in Fig. 4, 500- μm -thick hippocampal slices were used and the ACSF composition included 1 mM MgCl₂ and 2 mM CaCl₂.

Electrophysiological recordings

In most cases, we recorded from two slices simultaneously in the same recording chamber. We evoked synaptic responses by stimulating Schaffer collaterals with 0.2-ms pulses delivered using concentric bipolar stimulating electrodes (FHC), and recorded them extracellularly in CA1 stratum radiatum. Baseline responses were obtained by stimulating at 0.033 Hz using an intensity that yielded a half-maximal field potential slope. To induce LTP, four episodes of TBS were delivered at 0.1 Hz, using the same stimulation intensity as for baseline. TBS consists of 10 stimulus trains delivered at 5 Hz with each train consisting of four pulses at 100 Hz. Homosynaptic LTD was induced by delivering LFS (900 pulses at 1 Hz) at the same stimulation intensity as baseline²⁷. After the recording, slices were quickly frozen on dry ice for immunoblot analysis.

For two-pathway experiments, two stimulating electrodes were placed on either side of the recording field potential electrode to stimulate two independent synaptic inputs, alternating every 15 s. The absence of cross-pathway paired-pulse facilitation was used to ensure that the two inputs were independent of each other. KN-93 (Calbiochem) was dissolved in 100% DMSO (Sigma) and directly applied to the ACSF perfusion solution to reach a final concentration of 10 μ M in 0.1% DMSO.

Preparation of slices for SDS-PAGE

Homogenates of hippocampal slices were prepared by sonicating each slice on ice in 1 ml of resuspension buffer (10 mM sodium phosphate (pH 7.0), 100 mM NaCl, 10 mM sodium pyrophosphate, 50 mM NaF, 1 mM sodium orthovanadate, 5 mM EDTA, 5 mM EGTA, 1 μ M okadaic acid and 10 U ml⁻¹ aprotinin) for 20 s. The homogenates were centrifuged at 14,000g for 10 min at 4°C. The crude membrane pellets were resuspended in SDS sample buffer and loaded onto SDS-PAGE gels (7.5%). We usually loaded a control sample of hippocampus onto each gel at different protein concentrations to confirm that the protein concentration in our samples falls in the linear range of antibody detection. The resulting gels were transferred to PVDF membranes.

Immunoblot analysis

The phosphorylation-site-specific antibodies were generated and purified as described¹⁴. PVDF membranes were incubated in 25% methanol and 10% acetic acid (15 min). The membranes were then blocked with 1% BSA and 0.1% Tween-20 in PBS for 1 h, incubated for 1–2 h with the phosphorylation-site-specific antibodies (150–500 mg ml⁻¹ in blocking buffer), washed five times for 5 min with blocking buffer, and incubated for 1 h with alkaline phosphatase-conjugated anti-rabbit immunoglobulin (1:10,000 in blocking buffer). After final washes in blocking buffer (5 $\times$ 5 min) the membranes were immersed in chemiluminescence (ECF) substrate (Amersham) for 5–30 min. The membranes were dried between filter papers, and scanned in Storm 860 (Molecular Dynamics) at 750–900 V. The phosphorylation-site-specific antibodies were stripped from the membranes by incubation in 62.5 mM Tris (pH 6.8), 2% SDS and 0.7% 2-mercapto-ethanol at 60°C for 20 min; the membranes were then reprobed with anti-GluR1 C-terminal antibodies to estimate the total amount of GluR1. The scanned digital images were quantified using ImageQuant software (Molecular Dynamics). We analysed the relative amount of GluR1 phosphorylation by determining the ratio of the signals (P/C ratio) detected using the phosphorylation-site-specific antibodies (P) and the phosphorylation-independent C-terminal antibody (C). The P/C ratio of the control and the matching test slices was used for statistical analysis. For display purposes, this ratio was normalized to the control to derive the percentage of control values. The linearity of immunoblots was confirmed by analysing the ratio of the control samples loaded at different concentrations on each gel. The ECF method gives a better linear signal than conventional chemiluminescence (ECL) detection when looking at the ratio of signals.

Received 20 March; accepted 24 April 2000.

1. Bliss, T. V. P. & Collingridge, G. L. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* **361**, 31–39 (1993).
2. Bear, M. F. & Abraham, W. C. Long-term depression in hippocampus. *Annu. Rev. Neurosci.* **19**, 437–462 (1996).
3. Bear, M. F. A synaptic basis for memory storage in the cerebral cortex. *Proc. Natl Acad. Sci. USA* **93**, 13453–13459 (1996).
4. Malenka, R. C. & Nicoll, R. A. Long-term potentiation—a decade of progress? *Science* **285**, 1870–1874 (1999).
5. Barria, A., Muller, D., Derkach, V., Griffith, L. C. & Soderling, T. R. Regulatory phosphorylation of AMPA-type glutamate receptors by CaM-KII during long-term potentiation. *Science* **276**, 2042–2045 (1997).
6. Lee, H.-K., Kameyama, K., Huganir, R. L. & Bear, M. F. NMDA induces long-term synaptic depression and dephosphorylation of the GluR1 subunit of AMPA receptors in hippocampus. *Neuron* **21**, 1151–1162 (1998).
7. Kameyama, K., Lee, H.-K., Bear, M. F. & Huganir, R. L. Involvement of a postsynaptic protein kinase A substrate in the expression of homosynaptic long-term depression. *Neuron* **21**, 1163–1175 (1998).
8. Hollmann, M. & Heinemann, S. Cloned glutamate receptors. *Annu. Rev. Neurosci.* **17**, 31–108 (1994).
9. Seeburg, P. H. The molecular biology of mammalian glutamate receptor channels. *Trends Neurosci.* **16**, 359–365 (1993).
10. Squire, L. R. Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychol. Rev.* **99**, 195–231 (1992).
11. Zamanillo, D. *et al.* Importance of AMPA receptors for hippocampal synaptic plasticity but not for spatial learning. *Science* **284**, 1805–1811 (1999).
12. Roche, K. W., O'Brien, R. J., Mammen, A. L., Bernhardt, J. & Huganir, R. L. Characterization of multiple phosphorylation sites on the AMPA receptor GluR1 subunit. *Neuron* **16**, 1179–1188 (1996).
13. Barria, A., Derkach, V. & Soderling, T. Identification of the Ca²⁺/calmodulin-dependent protein kinase II regulatory phosphorylation site in the α -amino-3-hydroxyl-5-methyl-4-isoxazole-propionate type glutamate receptor. *J. Biol. Chem.* **272**, 32727–32730 (1997).
14. Mammen, A. L., Kameyama, K., Roche, K. W. & Huganir, R. L. Phosphorylation of the α -amino-3-hydroxy-5-methylisoxazole-4-propionic acid receptor GluR1 subunit by calcium/calmodulin-dependent kinase II. *J. Biol. Chem.* **272**, 32528–32533 (1997).
15. Derkach, V., Barria, A. & Soderling, T. R. Ca²⁺/calmodulin-kinase II enhances channel conductance of α -amino-3-hydroxy-5-methyl-4-isoxazolepropionate type glutamate receptors. *Proc. Natl Acad. Sci. USA* **96**, 3269–3274 (1999).
16. Banke, T. B. *et al.* Control of GluR1 AMPA receptor function by cAMP-dependent protein kinase. *J. Neurosci.* **20**, 89–102 (2000).
17. Mulkey, R. M., Herron, C. E. & Malenka, R. C. An essential role for protein phosphatases in hippocampal long-term depression. *Science* **261**, 1051–1055 (1993).

18. Mulkey, R. M., Endo, S., Shenolikar, S. & Malenka, R. C. Involvement of a calcineurin/inhibitor-1 phosphatase cascade in hippocampal long-term depression. *Nature* **369**, 486–488 (1994).
19. Malenka, R. C. *et al.* Long-term potentiation: an essential role for postsynaptic calmodulin and protein kinase activity. *Nature* **340**, 554–557 (1989).
20. Malinow, R., Schulman, H. & Tsien, R. W. Inhibition of postsynaptic PKC or CaMKII blocks induction but not expression of LTP. *Science* **245**, 862–866 (1989).
21. Lledo, P.-M. *et al.* Calcium/calmodulin-dependent kinase II and long-term potentiation enhance synaptic transmission by the same mechanism. *Proc. Natl Acad. Sci. USA* **92**, 11175–11179 (1995).
22. Otmakhov, N., Griffith, L. C. & Lisman, J. E. Postsynaptic inhibitors of calcium/calmodulin-dependent protein kinase type II block induction but not maintenance of pairing-induced long-term potentiation. *J. Neurosci.* **17**, 5357–5365 (1997).
23. Abeliovich, A. *et al.* Modified hippocampal long-term potentiation in PKC gamma-mutant mice. *Cell* **75**, 1253–1262 (1993).
24. Benke, T. A., Luthi, A., Isaac, J. T. & Collingridge, G. L. Modulation of AMPA receptor unitary conductance by synaptic activity. *Nature* **393**, 793–797 (1998).
25. Carroll, R. C., Lissing, D. V., von Zastrow, M., Nicoll, R. A. & Malenka, R. C. Rapid redistribution of glutamate receptors contributes to long-term depression in hippocampal cultures. *Nature Neurosci.* **2**, 454–460 (1999).
26. Shi, S. H. *et al.* Rapid spine delivery and redistribution of AMPA receptors after synaptic NMDA receptor activation. *Science* **284**, 1811–1816 (1999).
27. Dudek, S. M. & Bear, M. F. Homosynaptic long-term depression in area CA1 of hippocampus and the effects of NMDA receptor blockade. *Proc. Natl Acad. Sci. USA* **89**, 4363–4367 (1992).

Acknowledgements

We thank C. Doherty for help in preparing the antibodies and D. Bury for help with the manuscript. This work was supported by the Howard Hughes Medical Institute, NARSAD and The Grable Foundation.

Correspondence and requests for materials should be addressed to R.L.H. (e-mail: rhuganir@jhmi.edu).

Stable germline transformation of the malaria mosquito *Anopheles stephensi*

Flaminia Catteruccia*†, Tony Nolan*†, Thanasis G. Loukeris‡, Claudia Blass‡, Charalambos Savakis§, Fotis C. Kafatos‡ & Andrea Crisanti*

* Imperial College of Science, Technology and Medicine, Imperial College Road, London SW7 2AZ, UK

‡ European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany

§ Institute of Molecular Biology and Biotechnology, Research Centre of Crete, Foundation for Research and Technology Hellas, Heraklion, Crete, Greece

† These authors contributed equally to this work

Anopheline mosquito species are obligatory vectors for human malaria, an infectious disease that affects hundreds of millions of people living in tropical and subtropical countries. The lack of a suitable gene transfer technology for these mosquitoes has hampered the molecular genetic analysis of their physiology, including the molecular interactions between the vector and the malaria parasite. Here we show that a transposon, based on the *Minos* element¹ and bearing exogenous DNA, can integrate efficiently and stably into the germ line of the human malaria vector *Anopheles stephensi*, through a transposase-mediated process.

Genetic and genomic information on malaria vectors has expanded substantially during the last few years. However, further progress is hampered by the lack of germline transformation, a key technology in functional studies. In the fruitfly *Drosophila melanogaster* the development of gene transfer techniques has promoted, in a short time, a large flow of functional information on genes involved in embryogenesis, tissue modelling, intracellular signalling, neuronal organisation, behaviour and innate immunity. This success has prompted the development of methods for introducing

Table 1 Outcome of independent injection experiments

Exp	Embryos injected	Larvae	Adults	Crossing scheme	Fluorescent/total G ₁
I	167	46 (28%)	8 (5%)	wt outcross	0/3,694
II	158	27 (17%)	7 (4%)	wt outcross	0/3,165
III	123	29 (24%)	7 (6%)	intercross	0/813
IV	159	69 (43%)	27 (17%)	intercross	5/3,381
V	278	90 (32%)	42 (15%)	wt outcross	87/7,158
Total	885	261 (29%)	91 (10%)		92/18,211

A. stephensi embryos were injected with a mixture of pMinEGFP and pHSS6hSLMi20 in five different experiments (I–V). The percentages of embryos that hatched to larvae and survived to adulthood in each experiment are given in brackets. The total numbers of injected embryos, hatching larvae and surviving adults are also indicated. The surviving adults were crossed either among themselves (intercross) or to wild-type *A. stephensi* mosquitoes (wt outcross).

exogenous genes into the genome of insect pests of medical and agricultural importance. Success has only been achieved in the last five years, with the transformation of the mediterranean fruitfly *Ceratitis capitata*², the yellow fever mosquito *Aedes aegypti*^{3,4}, and the flour beetle *Tribolium castaneum*⁵.

We tested whether the *Minos* transposable element¹ from *Drosophila hydei* could integrate by active transposition into the germ line of the human malaria vector *Anopheles stephensi*. *Minos* was chosen on the basis of the findings that *Minos* transposase can mediate precise insertions in the genome of *Anopheles gambiae* cell lines and permits interplasmid transposition in *A. stephensi* embryos⁶. We have developed a plasmid vector, pMinEGFP, in which an enhanced version of the green fluorescent protein (EGFP)^{7,8} gene was placed within the inverted repeats of *Minos* under the control of the *actin5c* promoter from *D. melanogaster* (Fig. 1a). In five consecutive experiments (I–V) (Table 1), 885 *A. stephensi* embryos were injected with a mixture of pMinEGFP and the helper plasmid pHSS6hSLMi20⁹. The helper plasmid contains an intron-less version of the *Minos* transposase gene controlled by the heat shock protein 70 (*hsp70*) promoter of *D. melanogaster* and provides the enzyme for *Minos* transposition. The injections were performed on preblastoderm embryos following a standard procedure¹⁰ with an important modification. To prevent the hardening of the chorion, *A. stephensi* females were allowed to lay eggs in a 0.1-mM solution of p-nitrophenyl p'-guanidinobenzoate (pNpGB), an inhibitor of the enzyme phenoloxidase and thus of the first steps of the melanization process. In this solution, embryos remained soft and transparent for up to 3 h, during which time microinjections were performed.

An average of 29% of injected embryos survived and around 50% of the hatched larvae showed strong transient expression of EGFP, as monitored by fluorescence. Survival to adult stage (G₀) averaged 10% and was a good predictor of successful transformation (Table 1). In three early experiments (I–III), marked by low adult survival (average 5%), we detected no fluorescent individuals among 7,672 G₁ larvae derived from 22 G₀ survivors. In contrast, in two experiments (IV and V) that gave 16% adult survival, the progeny of 69 G₀ mosquitoes yielded 92 fluorescent individuals among the 10,539 G₁ larvae analysed. In experiment IV, adults were

Table 2 Transgenic families from experiment V

Breeding group	Founder	Fluorescent/total G ₁
9 males	A	0/779
9 males	B	14/1853
10 females	C	0/3277
14 females	D:	
	1,2	no eggs
	3–11	0/814
	12	14/166
	13	37/110
	14	22/159

The 42 surviving adults from experiment V (Table 1) were outcrossed to wild-type *A. stephensi* in four groups (A–D) of same-sex individuals. The 14 females from group D were allowed to lay eggs singly (D1–D14) to determine the number of single founders.

Table 3 Segregation of the EGFP marker in G₂ intercrosses

G ₂ family	–/–	+/–	+/+	χ^2
IV	65	135	47	4.76*
V _B	44	76	38	0.68*
V _{D12}	30	73	28	1.78*
V _{D13}	30	56	28	0.11*
V _{D14}	32	65	27	0.69*

Larvae derived from intercrosses between heterozygous G₂ individuals of the indicated families were scored for intensity of fluorescence. (+/+), larvae showing strong fluorescence, scored as homozygous; (+/–), larvae showing intermediate fluorescence, scored as heterozygous; –/–, larvae showing no fluorescence, scored as negative. *, Phenotypic distributions are consistent with the 1:2:1 ratio of genotypes as evaluated by χ^2 , d.f., 2.

intercrossed, so it was not possible to determine whether the five fluorescent G₁ larvae were derived from one or more founders (Table 1). In experiment V, the 42 G₀ survivors were split into four same-sex groups (A–D), which were outcrossed separately (Table 2). Two of these groups (A, C) did not produce fluorescent progeny. The G₀ males of group B were tested collectively and it was not determined whether the observed positive G₁ larvae had originated from one or more founders. The 14 G₀ females of group D were group-mated and laid eggs in isolation. Among these mosquitoes, three different founders (V_{D12}, V_{D13} and V_{D14}) generated fluorescent progeny. In summary, the 92 fluorescent G₁ individuals from experiments IV and V were derived from a minimum of five independent G₀ founders, representing a transformation frequency of 7% (5/69 surviving adults). This frequency is higher than that reported in *D. melanogaster* and *C. capitata* transformation experiments using *Minos* marked with the *white* gene marker^{2,11}, but is comparable to that achieved using EGFP as a selection marker (C.S., personal observation).

Fluorescent positive and negative G₂ larvae generated by all the G₁ founders were recorded in nearly equal numbers, in agreement with the 1:1 ratio expected from mendelian segregation of a single transgene insertion. Positive heterozygous G₂ adults were then intercrossed to obtain homozygous lines. We attributed an additional phenotype of stronger fluorescence in the G₃ generations to the homozygous genotype. The three fluorescent phenotypes (strong, weak and negative, Fig. 1b) were represented in approximately the 1:2:1 ratio expected from mendelian segregation of a single transgene (Table 3). To determine the nature of the integration events, we performed Southern blotting of transformed mosquitoes with probes spanning the arms of *Minos* (probe M), the EGFP sequence (probe EGFP) or the ampicillin-resistance gene present in the backbone of plasmid pMinEGFP (probe AMP) (Fig. 1c). In both *Hind*II and *Eco*RI digests from five representative families, two bands of variable size hybridized with probe M, in agreement with the occurrence of a single integration event. The size variation of the bands was consistent with single integrations occurring at different chromosomal sites (Fig. 1c). The integrity of the internal part of the transposon was confirmed by the presence of a single band of constant size that hybridized to the EGFP probe

Table 4 Sequence of integration sites

Family	Flanking sequence	<i>Minos</i> arm
IV	CGTATCGAGGCAAGTGT TA cgagccccaacc	left
V _{D12}	gggtggggctcg TAG TAGTGAAGCAGCCCC	right
V _{D14}	AAATCTATTGCGCCCTG TA cgagccccaacc	left
V _{D14}	gggtggggctcg TATA AGGTGACATTAGTAA	right

Clones containing putative *Minos* insertions were isolated from *Eco*RI Lambda Zap libraries of families IV, V_{D12} and V_{D14} and sequenced. The unique *A. stephensi* flanking sequences are shown in capitals, *Minos* left and right arm sequences are shown in lower case and the TA dinucleotide of the insertion site is shown in bold. The *Minos* arm cloned (left or right) is also indicated. Note that both *Minos* arms of the line V_{D14} were cloned and sequenced, confirming the absolute precision of the transposition event. The *A. stephensi* flanking sequences were deposited in the EMBL Nucleotide Sequence Database with accession numbers AJ277720 (line IV), AJ277721 (line V_{D12}), AJ277722 (line V_{D14} left arm) and AJ277723 (line V_{D14} right arm).

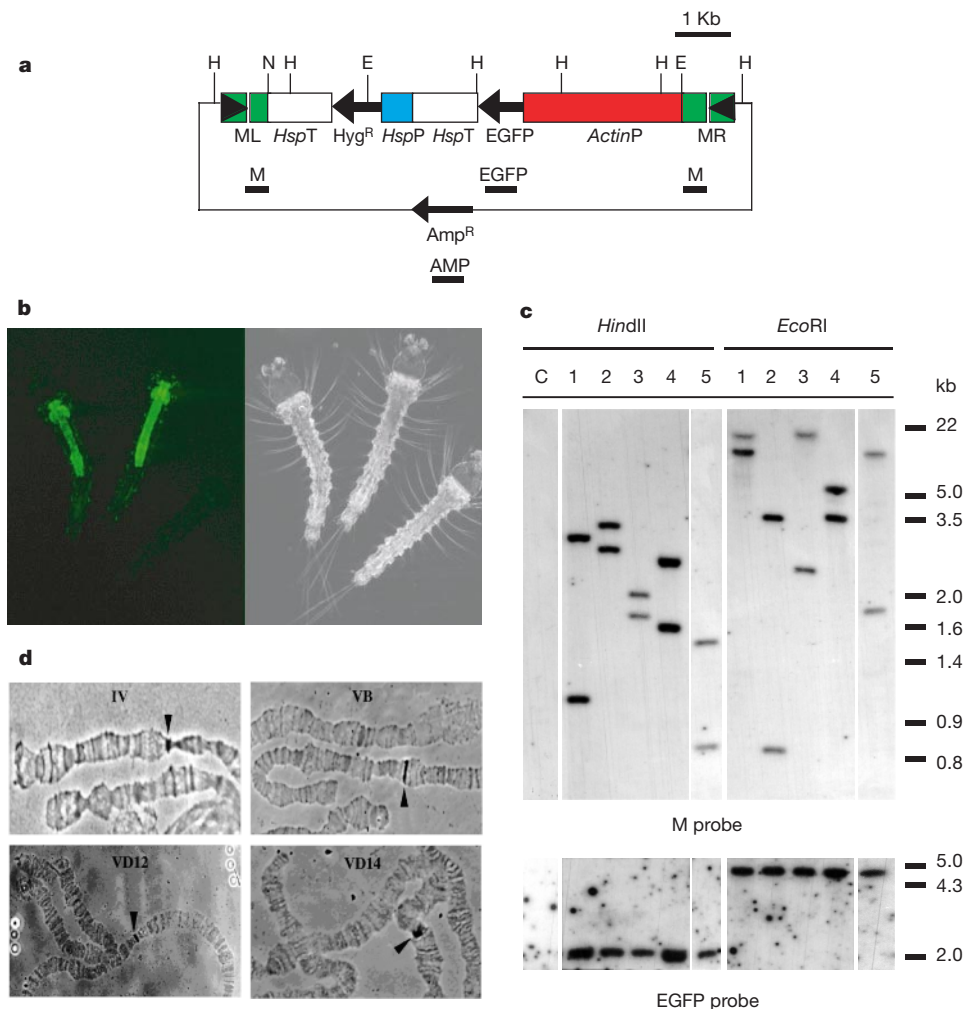


Figure 1 The Minos element integrates into the *A. stephensi* genome. **a**, Map of the transformation vector pMinEGFP. *actinP*, *D. melanogaster actin5c* promoter; *hspP*, *D. melanogaster hsp70* promoter. *hspT*, *D. melanogaster hsp70* terminator sequence. The EGFP, ampicillin-resistance (Amp^R) and hygromycin-resistance (Hyg^R) genes are indicated by black arrows and the left (ML) and right (MR) arms of *Minos* are in green, with the inverted repeats represented by black triangles. H, *Hind*III; E, *Eco*RI; N, *Not*I. Black bars represent the three probes (M, EGFP and AMP) used in Southern blot experiments. **b**, Confocal fluorescence and transmission microphotographs of putative heterozygous

(left) and homozygous (middle) transgenic larvae expressing EGFP compared with a wild-type mosquito (right). **c**, Southern blotting of genomic DNA from transgenic families IV, V_{D12} , V_{D13} , V_{D14} and V_B (lanes 1–5 respectively) digested with either *Hind*III or *Eco*RI. Lane C, DNA from wild-type *A. stephensi* mosquitoes, digested with *Hind*III. Upper panel, hybridization with probe M; lower panel, hybridization with probe EGFP (Fig. 1a). **d**, *In situ* hybridization on the polytene chromosomes of ovarian nurse cells of transformed lines IV, V_B , V_{D12} and V_{D14} . Arrowheads, distinct chromosomal insertions of *Minos* in the different lines.

(Fig. 1c). The AMP probe failed to produce a signal, as expected for *Minos*-mediated integration (data not shown).

To verify the precision of transposition, we determined the sequences of the junctions between the inserted element and the *A. stephensi* genome. Genomic DNA libraries from families IV, V_{D12} and V_{D14} were screened with probe M. All positive clones were sequenced and the presence of precise *Minos*-mediated integrations in these three families was confirmed (Table 4). Each of the cloned insertion arms included a complete *Minos* inverted repeat, a TA dinucleotide and an additional sequence, distinct for each case and presumably derived from a different site of the *A. stephensi* genome. We then performed *in situ* hybridizations with plasmid pMinEGFP on polytene chromosome preparations obtained from four representative families. In each case, a single insertion site was detected per chromosome complement, and this site was distinct in each family, confirming the results of the Southern blot analysis (Fig. 1d).

Our results show that a *Minos* transposon carrying exogenous DNA can be integrated by a transposase-mediated mechanism into the genome of *A. stephensi* mosquitoes at high frequency. Furthermore, we have demonstrated that EGFP is a reliable selectable

marker for anopheline transformation, as for other insect species^{5,12,13}. This is the first reliable system for germline transformation of an anopheline vector of human malaria. We expect this technology to be successfully extended to the most important malarial vector, *A. gambiae*, as cultured cells of this mosquito species can be transformed in a transposase-mediated manner using a very similar *Minos* transposon⁶. The availability of this gene transfer technique may contribute to progress in our understanding of the mechanisms of vector competence, and could open the prospect of engineering anopheline mosquitoes that are refractory to *Plasmodium*. □

Methods

Plasmid construction

The pMinEGFP transformation vector was derived from plasmid pMinNot² by inserting two DNA fragments: an *Eco*RI–*Not*I fragment containing the EGFP gene (Clontech) under the control of the *actin5c* promoter from *D. melanogaster* and a *Not*I–*Not*I fragment containing the hygromycin B phosphotransferase gene (*hyg*) controlled by the *hsp70* promoter from *D. melanogaster*. The *hyg* gene was introduced as an additional selectable marker in the event that selection with EGFP could not be achieved. Attempts to verify

whether *hyg* could function as a selectable marker did not give reproducible results. The helper plasmid pHSS6hsILMi20 has been described⁹.

Embryo microinjection

Blood-fed *A. stephensi* mosquitoes were allowed to lay eggs 48–72 h after a blood meal. Eggs were laid in a solution of 0.1 mM p-nitrophenyl p'-guanidinobenzoate (pNpGB) (Sigma) in isotonic buffer (150 mM NaCl, 5 mM KCl, 10 mM HEPES; 2.5 mM CaCl₂; pH 7.2). Eggs were then left in pNpGB solution until injection, which was carried out 90–120 min after oviposition. The embryos were microinjected by using glass needles (Eppendorf) with a mixture of the helper plasmid pHSS6hsILMi20 (100 µg ml⁻¹) and the pMinEGFP plasmid (400 µg ml⁻¹) in injection buffer (5 mM KCl, 0.1 mM sodium phosphate, pH 6.8). After injection, the embryos were placed in isotonic buffer and transferred back to an insectary (80% humidity; 26 °C) to hatch. Injected embryos were not subjected to heat shock as previous experiments have shown that the baseline activity of *hsp70* promoter provided enough *Minos* transposase to mediate integration into the genome of *A. gambiae* cells⁶. Hatched larvae were analysed on an inverted microscope at a wavelength of 490 nm to detect EGFP expression.

Southern blot analyses and sequencing of integration sites

Genomic DNA from transformed adult or fourth instar larvae mosquitoes (G₃ generation) and from wild-type adult mosquitoes was digested with the restriction endonucleases *Hind*II or *Eco*RI. Digested genomic DNA (~4 µg per lane) was separated on a 0.8% agarose gel and transferred onto a nylon membrane. The membrane was hybridized overnight at 65 °C with three ³²P-labelled probes: (1) probe M, a polymerase chain reaction (PCR) product encompassing sequences of the left and right arms of *Minos*; (2) probe EGFP, spanning the EGFP gene; and (3) probe AMP, encompassing sequences of the ampicillin gene present in the backbone of plasmid pMinEGFP. Between hybridizing with each probe, the membrane was washed with a boiling 10-mM EDTA/0.1% SDS solution for 20 min to dehybridize it. Genomic Lambda Zap *Eco*RI libraries (Stratagene) were constructed from the DNA extracted from G₃ larvae of transgenic lines IV, V_{D12} and V_{D14} and hybridized with probe M. The cloned insertion sites were sequenced using primers annealing internally to the inverted repeats (pMR, 5'-CGAGTTAAATGCGTAATGC-3' and pML, 5'-GCTCTTCTTGAGATTAAGG-3').

In situ hybridization

The integration of the *Minos* element in the mosquito genome was visualized by *in situ* hybridization experiments performed on the polytene chromosomes of ovarian nurse cells prepared from four out of the five established lines as described¹⁴. The whole plasmid pMinEGFP was used as a template to produce a biotinylated probe.

Received 24 March; accepted 4 May 2000.

1. Franz, G. & Savakis, C. *Minos*, a new transposable element from *Drosophila hydei*, is a member of the *Tc1*-like family of transposons. *Nucleic Acids Res.* **19**, 6646 (1991).
2. Loukeris, T. G., Livadaras, I., Arca, B., Zabalou, S. & Savakis, C. Gene transfer into the medfly, *Ceratitis capitata*, with a *Drosophila hydei* transposable element. *Science* **270**, 2002–2005 (1995).
3. Coates, C. J., Jasinskiene, N., Miyashiro, L. & James, A. A. *Mariner* transposition and transformation of the yellow fever mosquito, *Aedes aegypti*. *Proc. Natl Acad. Sci. USA* **95**, 3748–3751 (1998).
4. Jasinskiene, N. *et al.* Stable transformation of the yellow fever mosquito, *Aedes aegypti*, with the *Hermes* element from the housefly. *Proc. Natl Acad. Sci. USA* **95**, 3743–3747 (1998).
5. Berghammer, A. J., Klingler, M. & Wimmer, E. A. A universal marker for transgenic insects. *Nature* **402**, 370–371 (1999).
6. Catteruccia, F. *et al.* Toward *Anopheles* transformation: *Minos* element activity in anopheline cells and embryos. *Proc. Natl Acad. Sci. USA* **97**, 2157–2162 (2000).
7. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805 (1994).
8. Cormack, B. P., Valdivia, R. H. & Falkow, S. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**, 33–38 (1996).
9. Klinakis, A. G., Loukeris, T. G., Pavlopoulos, A. & Savakis, C. Mobility assays confirm the broad host range activity of the *Minos* transposable element and validate new transformation tools. *Insect Mol. Biol.* (in the press).
10. Miller, L. H. *et al.* Stable integration and expression of a bacterial gene in the mosquito *Anopheles gambiae*. *Science* **237**, 779–781 (1987).
11. Loukeris, T. G., Arca, B., Livadaras, I., Dialektaki, G. & Savakis, C. Introduction of the transposable element *Minos* into the germ line of *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **92**, 9485–9489 (1995).
12. Pinkerton, A. C., Michel, K., O'Brochta, D. A. & Atkinson, P. W. Green fluorescent protein as a genetic marker in transgenic *Aedes aegypti*. *Insect Mol. Biol.* **9**, 1–10 (2000).
13. Tamura, T. *et al.* Germline transformation of the silkworm *Bombyx mori* L. using a piggyBac transposon-derived vector. *Nature Biotechnol.* **18**, 81–84 (2000).
14. Kumar, V. & Collins, F. H. A technique for nucleic acid *in situ* hybridization to polytene chromosomes of mosquitoes in the *Anopheles gambiae* complex. *Insect Mol. Biol.* **3**, 41–47 (1994).

Acknowledgements

We thank A. Richman at EMBL for the initial efforts for anopheline transformation and D. Prager and the MacArthur Vector Biology Network for the collegial interactions that greatly stimulated the project. The research was supported by a Network grant of the training and Mobility Program of the European Community, by Implex Ltd, by WHO-TDR, by the John D. and Catherine T. MacArthur Foundation, the SFB 544 of the Deutsche Forschungsgemeinschaft and an individual fellowship of the Biotechnology Program of the EU (T.G.L.).

Correspondence and requests for materials should be addressed to A.C.

A small-molecule nitroimidazopyran drug candidate for the treatment of tuberculosis

C. Kendall Stover*, Paul Warren*, Donald R. VanDevanter*, David R. Sherman†, Taraq M. Arain*, Michael H. Langhorne*, Scott W. Anderson*, J. Andrew Towell*, Ying Yuan*, David N. McMurray‡, Barry N. Kreiswirth§, Clifton E. Barry|| & William R. Baker*

* PathoGenesis Corporation, 201 Elliott Avenue West, Seattle, Washington 98119, USA

† Department of Pathobiology, School of Public Health and Community Medicine, University of Washington, Seattle, Washington 98195, USA

‡ Department of Medical Microbiology and Immunology, Texas A&M University Health Science Center, Room 463 Reynolds Medical Building College Station, Texas 77843, USA

§ Tuberculosis Center, Public Health Research Institute, New York, New York 10016, USA

|| Tuberculosis Research Section, Laboratory of Host Defenses, NIAID, NIH, Twinbrook 2, Room 239, 12441 Parklawn Drive, Rockville, Maryland 20852, USA

Mycobacterium tuberculosis, which causes tuberculosis, is the greatest single infectious cause of mortality worldwide, killing roughly two million people annually¹. Estimates indicate that one-third of the world population is infected with latent *M. tuberculosis*². The synergy between tuberculosis and the AIDS epidemic^{3–5}, and the surge of multidrug-resistant clinical isolates of *M. tuberculosis* have reaffirmed tuberculosis as a primary public health threat. However, new antitubercular drugs with new mechanisms of action have not been developed in over thirty years. Here we report a series of compounds containing a nitroimidazopyran nucleus that possess antitubercular activity. After activation by a mechanism dependent on *M. tuberculosis* F420 cofactor, nitroimidazopyrans inhibited the synthesis of protein and cell wall lipid. In contrast to current antitubercular drugs, nitroimidazopyrans exhibited bactericidal activity against both replicating and static *M. tuberculosis*. Lead compound PA-824 showed potent bactericidal activity against multidrug-resistant *M. tuberculosis* and promising oral activity in animal infection models. We conclude that nitroimidazopyrans offer the practical qualities of a small molecule with the potential for the treatment of tuberculosis.

A series of bicyclic nitroimidazofurans, originally investigated as radiosensitizers for use in cancer chemotherapy⁶, were found to possess activity against cultured replicating *M. tuberculosis* (MTB) and had significant *in vivo* activity in a murine infection model^{7–9}. The lead compound in this series, CGI-17341 was mutagenic, discouraging further investigation of the antibacterial activity of the compound series⁹. These studies suggested, however, that the bicyclic nitroimidazoles might be potential antitubercular agents. A series of 328 3-substituted nitroimidazopyrans (NAPs) were synthesized on the basis of the structure of CGI-17341 (Fig. 1; W.R.B. *et al.*, unpublished data). Over 100 of these compounds possessed

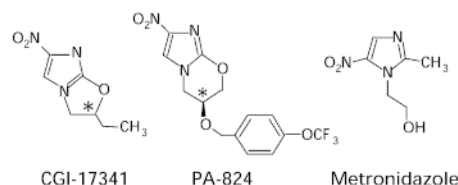


Figure 1 Structure of Metronidazole, CGI-17341 and PA-824. Asterisk denotes the C3 position and NAP chiral centre.

substantial antitubercular activity, matching or exceeding that of CGI-17341. The active compounds lacked the mutagenicity previously associated with bicyclic nitroimidazoles (data not shown). Structure activity relationship studies focusing on antitubercular activity revealed substantial variety in the tolerated substituents at C3, but optimal activity was achieved with lipophilic groups. The stereochemistry at C3 was important for activity, as the *S* enantiomers were generally at least 10-fold more active than the *R* enantiomers. NAP activity was found to be highly specific for the MTB complex, and showed only modest or no activity against mycobacteria outside the MTB complex (*M. avium*, *M. smegmatis*, *M. chelonae* and *M. fortuitum*).

One NAP compound, PA-824, exhibited a sub-micromolar minimal inhibitory concentration (MIC) against MTB (Fig. 1). The MIC values of PA-824 against a panel of MTB pan-sensitive and rifampin mono-resistant clinical isolates ranged from 0.015 to 0.25 $\mu\text{g ml}^{-1}$ (Table 1(a)). Poly- and multidrug resistant strains of MTB exhibited comparable susceptibility to PA-824, indicating that there is no cross-resistance with current MTB drugs. (Table 1(b)). PA-824 was bactericidal against MTB by post-treatment dilution at levels only fourfold greater than its MIC (data not shown). We also investigated NAP activity against non-replicating MTB in an anaerobic culture model because metronidazole, a structurally related antibiotic used to treat anaerobic infections, possesses activity against static MTB surviving under anaerobic conditions¹⁰. After adaptation to microaerophilic culture, MTB does not replicate; thus, any drug activity under these conditions would suggest a mode of killing that is independent of cell growth. PA-824 exhibited bactericidal activity comparable to that of the related metronidazole, whereas parent compound CGI-17341 and the front-line tuberculosis drug isoniazid (INH) were much less active against non-replicating MTB in this model (Fig. 2).

Over 50 NAP compounds with demonstrated activity against cultured MTB were screened for *in vivo* antitubercular activity using a short-term murine model of infection with an MTB reporter strain expressing firefly luciferase (rMTB-*lux*) (data not shown)¹¹. Although PA-824 was not the most potent NAP against cultured MTB clinical isolates, it was the most active in infected mice when orally administered at 25 mg kg^{-1} . This indicated that PA-824 might possess more desirable pharmacokinetic properties than the other more potent NAP compounds that we tested. Further studies in mice at 25, 50 and 100 mg kg^{-1} PA-824 daily for 10 days resulted in reductions of mycobacterial burden in both spleen and lung tissues that were comparable to that of INH at 25 mg kg^{-1} (Fig. 3a). Studies with PA-824 at doses of 50 or 100 $\text{mg kg}^{-1}\text{day}^{-1}$ for longer durations also resulted in significant reductions of MTB burden in murine

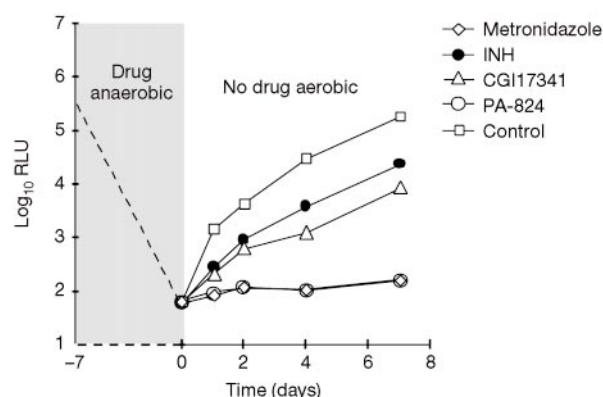


Figure 2 NAP activity on static MTB maintained in microaerophilic culture. Log-phase rMTB-*lux* bacilli were mixed with the indicated drugs (10 $\mu\text{g ml}^{-1}$) and placed in loose-capped tubes inside an activated Gaspak chamber (Difco). After 7 days of microaerophilic incubation at 37 °C, cells were pelleted by centrifugation (1,800g, 10 min), resuspended in fresh aerated, drug-free medium, and recovery was monitored by determining relative light units (RLU)²⁴. The reduction in CFU compared with that of untreated control cultures was also determined by plating at days 0 and 4 to evaluate bactericidal activity. Log₁₀ CFU reductions for PA-824, metronidazole, CGI-17341 and INH were 3.5, 2.8, 1.8 and 1.4 at day 0, respectively, and 3.1, 3.2, 1.5 and 1.0 at day 4, respectively.

lungs, comparable to INH at 25 $\text{mg kg}^{-1}\text{day}^{-1}$ (Fig. 3b). In an MTB aerosol challenge guinea pig model, oral administration of PA-824 at 40 mg kg^{-1} once daily for 30 days resulted in statistically significant reductions of MTB bacilli in lungs and spleens, comparable to the effect of INH when compared to controls (Fig. 3c). The therapeutic activity observed for PA-824 in murine and guinea pig infection models was also well under the acute and chronic toxic thresholds for PA-824 in mice. Toxic thresholds were determined to be greater than 1,000 mg kg^{-1} single dose (acute) and greater than 500 mg kg^{-1} when administered daily for 28 days (chronic). In summary, these data from *in vivo* animal infection models indicate that there is significant potential for safe oral delivery of PA-824 in the treatment of tuberculosis infection.

The activity of PA-824 against both multidrug-resistant and non-replicating microaerophilic MTB suggested that the NAPs act by a new mechanism. Studies on the macromolecular effects of PA-824 on MTB revealed that both protein and lipid synthesis were substantially inhibited (Fig. 4a) at drug concentrations and time intervals when nucleic acid synthesis was unaffected (data not shown). Lipid synthesis was not affected by treatment with amika-

Table 1 PA-824 drug susceptibilities for MTB clinical isolates

(a)			(b)		
Strain	MIC ($\mu\text{g ml}^{-1}$)		Strain	Drug resistance	PA-824 MIC ($\mu\text{g ml}^{-1}$)
	INH	PA-824			
H ₃₇ Rv	0.03	0.13	TN565	R,S,EM,ET,K,Cl	0.06
TN675*	0.03	0.015	TN576	I,R,S,EM,ET,K	0.06
TN913	0.03	0.06	TN702	I,S,EM,P	0.25
TN994*	0.03	0.03	TN715	I,R,EM,P	0.06
TN1008	0.06	0.06	TN768	I,R,S,EM,ET	0.06
TN1037	0.03	0.06	TN772	I,R,EM	0.03
TN1040	0.03	0.13	TN1195	I,S,EM	0.25
TN1051	0.03	0.06	TN1314	I,R	0.03
TN1082	0.03	0.06	TN1618	I,R,S,EM,ET,Cl	0.06
TN2351	0.06	0.06	TN1811	I,R,S,EM	0.03
TN2524	0.06	0.25	TN2557	I,R,S,EM,CA	0.13
TN3183	0.03	0.015			
TN3979	0.06	0.13			
TN4259	0.03	0.06			

(a) PA-824 and isoniazid drug susceptibilities were determined on 12 pan-susceptible and 2 mono-rifampin resistant (asterisk) clinical isolates. The PA824 susceptibilities were also tested on 9 multidrug-resistant and 2 poly-resistant MTB strains.

(b) Twenty of the twenty-five sensitive and resistant clinical isolates tested were previously determined to be genetically distinct by IS6110 genotyping. I, isoniazid; R, rifampin; S, streptomycin; EM, ethambutol; ET, ethionamide; K, kanamycin; P, pyrazinamide; Cl, ciprofloxacin; CA, capreomycin.

cin, an aminoglycoside protein synthesis inhibitor, indicating that the effect on lipid synthesis is specific and not a consequence protein synthesis inhibition in MTB (data not shown). Further analysis of the effects on cell wall lipids indicated that PA-824 produced an accumulation of hydroxymycolic acid with a concomitant reduction in ketomycolate (Fig. 4a, b). As hydroxymycolate is a known biosynthetic precursor of cell wall ketomycolate¹², PA-824 may inhibit an enzyme or deplete a cofactor responsible for the oxidation of hydroxymycolate to ketomycolate.

Metronidazole requires activation by nitroreduction for activity against anaerobic bacteria¹³. To examine whether the nitroimidazole nucleus of PA-824 underwent a similar reductive activation, we administered radiolabelled PA-824 to the susceptible *M. tuberculosis* complex variant BCG and a spontaneous NAP-resistant isolate, SM3, with a greater than 1,000-fold increase in MIC. Susceptible organisms converted PA-824 into an assortment of more polar metabolites, whereas resistant organisms did not effect any apparent transformations (Fig. 5, lanes 1 and 2). The SM3 strain was also found to be less proficient at reducing other nitroaromatic compounds such as 8-nitroquinolone (data not shown). Together, these data indicate that the metabolic activation of NAP and PA-824 by MTB may involve a nitro-reduction step analogous to that required for metronidazole activation¹⁴, and that the resistant mutants had lost their ability to carry out nitro-reduction on PA-824.

To identify the genes or structural elements involved in the metabolism and activation of PA-824, we carried out subtractive enrichment¹⁵ using complementary DNA derived from PA-824-

susceptible and resistant BCG messenger RNA. Sequence analysis of cDNA difference products and subsequent gene complementation studies revealed that only one open reading frame, Rv0407, encoding a protein with 89% identity to an F420-dependent glucose-6-phosphate dehydrogenase (Fgd) of *Mycobacterium smegmatis*^{16–18} was necessary and sufficient to restore NAP susceptibility to these BCG mutants. Complementation with Fgd also restored metabolic activation of radiolabelled PA-824 (Fig. 5, lane 3). Additional analysis of both BCG and MTB NAP-resistant mutants revealed a variety of point mutations in the identical BCG and MTB *fgd* genes that result in either missense substitutions or premature polypeptide termination. We also identified NAP-resistant mutants that lack lesions in the *fgd* gene, indicating that uncharacterized mutations in other genetic loci may also confer NAP resistance. The total frequency of resistance by any mechanism to PA-824 was determined by fluctuation analysis in MTB strain H37Rv to be 9.0×10^{-7} , slightly less than that of 1.3×10^{-6} or INH.

The NAPs represent a class of antitubercular compounds that act by a new mechanism, but that share some interesting parallels and contrasts with INH. Like INH, the NAP PA-824 pro-drug requires bacterial activation, albeit by a different F420-dependant mechanism. PA-824 also inhibits a step in the synthesis of cell wall mycolates. PA-824 inhibits a more terminal step than INH, however, namely the oxidation of hydroxymycolates to ketomycolates, a lipid class making up the mycobacterial pseudo-outer membrane and over one-third of the dry weight of MTB¹⁹. Whereas ketomycolate appears to be essential for the normal growth and survival of MTB

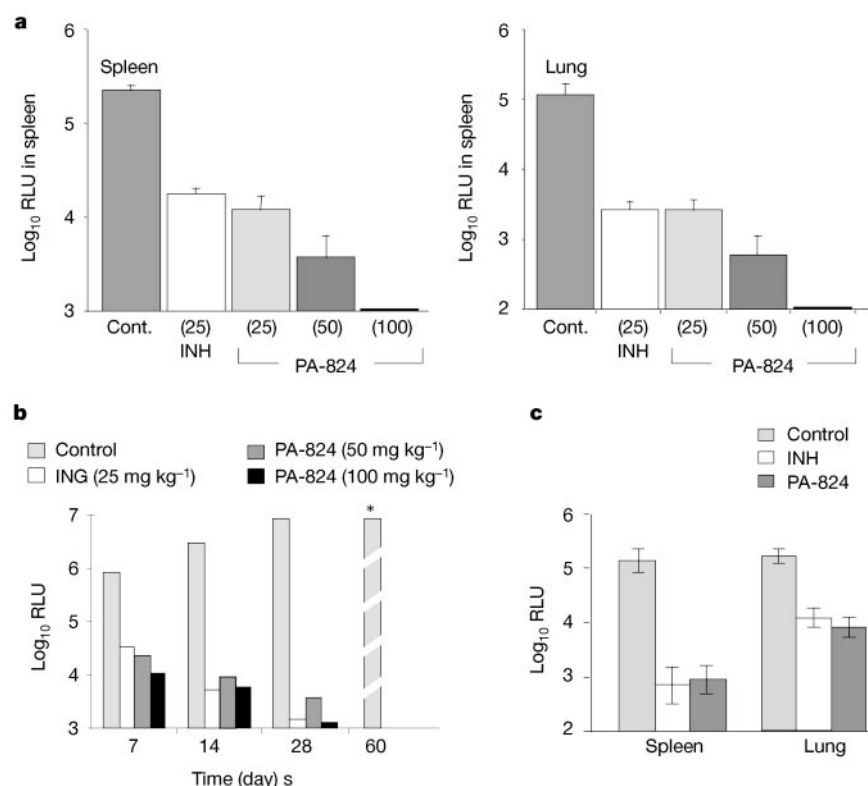


Figure 3 *In vivo* activity of PA-824 in murine and guinea pig models of tuberculosis infection. **a**, BALB/c mice were infected intravenously with roughly 10^6 CFU of an MTB H37Rv reporter strain for 4 days before daily oral administration of drug at indicated levels for 10 days. RLU from surviving MTB reporter strain in mouse lungs were determined 10 days after treatment initiation¹¹. **b**, BALB/c mice were infected intravenously with roughly 10^6 CFU of an MTB H37Rv reporter strain for 4 days before daily oral administration of drug at the indicated levels for 10 days. RLU from surviving MTB reporter organisms in mouse lungs were determined at indicated intervals of 7, 14, 28 and 60 days after treatment initiation¹¹. Asterisk, untreated control mice were killed before day 60 because of moribundity owing to MTB infection, and day 28 control RLU and CFU values are used

for day 60. All treated animals survived beyond day 60 but did not exhibit an RLU-MTB burden above background values. CFU reductions in drug-treated animals at day 60 were also compared with last surviving untreated controls animals at day 28. Day 60 Log₁₀ CFU reductions for INH at 25 mg kg⁻¹, and PA824 at 50 mg kg⁻¹ and 100 mg kg⁻¹ were roughly 5.6, 4.6 and 4.8 logs below day 28 untreated animals. **c**, Guinea pigs were infected with roughly 10 CFU of virulent *M. tuberculosis* (strain H37RV) by aerosol. After a 30-day incubation period, experimental animals were treated daily with 25 mg kg⁻¹ of INH or 40 mg kg⁻¹ of PA-824 per oral for 30 days. MTB CFUs were determined in lung and spleen tissue 30 days after treatment (60 days post infection). Group sizes for control, INH and PA-824 were 4, 6 and 6 guinea pigs, respectively.

in macrophages²⁰, NAP lethality may lie in the concurrent effect on protein synthesis, but this possibility remains to be investigated in detail. As demonstrated by PA-824 culture and models of animal infection, the NAPs possess bactericidal activity against MTB comparable to that of INH at doses that are safe in mice and guinea pigs. PA-824 and the NAPs are also small, orally active molecules that should be amenable to large-scale production, making them practical and attractive candidates for the clinical development of a new antitubercular therapy. In addition, activity against multidrug-resistant tuberculosis and demonstrable activity by PA-824 against static MTB could offer significant potential for

reducing the duration of therapy over current antitubercular therapies that target only growing organisms. □

Methods

Culture conditions

Cultures of the recombinant *M. bovis* BCG or MTB were grown from -70°C frozen stocks to log phase in Middlebrook 7H9 broth (Difco Laboratories or Remel) and refrigerated at 4°C as individual aliquots until needed. Inoculum for the rBCG-*lux* assay was prepared from the refrigerated stock and incubated at 37°C (rolling) to an absorbance at 540 nm of about 0.1. For drug susceptibility testing of static mycobacteria persisting under micro-aerophilic conditions, we first prepared cultures in Middlebrook 7H9 broth (Difco) and supplemented them with 10% ADC (prepared in-house), 0.5% Tween 80 (Difco) and 0.5% glycerol (Sigma).

Drug preparation

Isoniazid (INH), streptomycin (SM), Ethambutol (EMB), amikacin (AMK) and metronidazole (MZ) were obtained from Sigma and were prepared in sterile deionized water. NAP experimental compounds were prepared in dimethyl sulphoxide (Sigma) for *in vitro* drug susceptibility testing.

Drug susceptibility testing

Aerobic minimal inhibitory concentrations (MIC) were determined for BCG and MTB strains by BACTEC radiometric analysis as described²¹. Colony-forming-unit (CFU) determinations were made on Middlebrook 7H11 agar medium using an Autoplate spiral plater to generate dilutions (Spiral Biotech). After 2–3 weeks incubation at 37°C , colonies were counted using a Laser Colony Scanner (Spiral Biotech). Bioluminescence of firefly luciferase expressed by viable recombinant MTB-*lux* or BCG-*lux* was determined as described²² using either an Autolumat model 96P plate luminometer or an Autolumat model 953B tube luminometer (Wallac Instruments). Clinical isolates used in drug susceptibility testing were strain typed by IS6110 analysis as described²³.

Static microaerophilic rMTB-*lux* assay

Log-phase cultures of rMTB-*lux*, consisting of strain H37Rv containing integrated vector pMV361-*lux*¹¹, cultures were transferred in 5-ml aliquots into 15-ml screw-capped centrifuge tubes, and sonicated in a Vibra-Cell cup horn sonicator (Sonics & Materials) for 30 s at 20% amplitude, 0.1-s pulses. INH, CGI-17341, PA-824 or metronidazole was added to individual tubes to a final concentration of $10\text{ }\mu\text{g ml}^{-1}$. Tubes were loosely capped to allow gas circulation and placed in a rack inside an anaerobic jar. Anaerobic conditions were established in Gaspak anaerobic jars (BBL), using Anaerogen gas generation envelopes (Unipath), and monitored with anaerobic indicator strips (BBL). The jar was incubated for 7 days at 37°C . The static cultures were collected by low-speed centrifugation, resuspended in drug-free media, mildly sonicated to disperse clumped mycobacteria into single-cell suspensions and incubated at 37°C with rotating aeration. CFU and relative light units (RLU) were monitored until recovery of RLU to day 0 control level.

Isolation of NAP resistant mutants

BCG (substrain Montreal) was cultured in Middlebrook 7H9 media to mid log phase. We plated $150\text{-}\mu\text{l}$ aliquots on Middlebrook 7H10 agar containing $5\text{ }\mu\text{g ml}^{-1}$ of PA-824 and incubated them at 37°C for 2 weeks. Spontaneous PA-824 resistant colonies were observed at a frequency of 10^{-6} and isolated at random for further study.

Labelling and analysis of protein and lipids

For protein or lipid labelling, 5 ml log-phase rolling cultures of BCG (absorbance₅₄₀ = 0.2) in 7H9 broth plus ADC enrichment were treated with PA-824 at the indicated concentration for 3 h at 37°C . Control or drug-treated cultures were adjusted to $20\text{ }\mu\text{Ci ml}^{-1}$ ³⁵S-protein-labelling mix (Expre³⁵S³⁵S, NEN) and incubated an additional 60 min at 37°C . Cells were chilled for 10 min on ice, collected by centrifugation (5 min at 2000g), washed once with cold PBS, resuspended in 0.5 ml water containing $5\text{ }\mu\text{g ml}^{-1}$ leupeptin and transferred to a 2-ml screw-cap tube containing 0.75 g of 0.1-mm glass beads. Cells were disrupted by $3\times 20\text{-s}$ cycles in a Bio101 FP120 bead-beater, with power set at 6.5, and kept on ice between cycles. Aliquots of lysate were subjected to TCA-precipitation and incorporated radioactivity quantitated by liquid scintillation counting. For lipid labelling, control and drug-treated cultures were adjusted to $2\text{ }\mu\text{Ci ml}^{-1}$ ¹⁴C-acetate (NEN) and incubated for 15 min at 37°C . Cells were collected and washed as above. Cell pellets were resuspended in 1 ml 30% (w/v) KOH in 1:1 ethanol:H₂O, transferred to glass tubes and incubated 16 h at 70°C . Samples were cooled on ice and 0.5 ml of concentrated HCl added. Samples were extracted three times with diethyl ether, and the combined diethyl ether extracts evaporated to dryness. Samples were dissolved in 0.2 ml 2:1 chloroform:methanol. Labelled lipid extract ($10\text{ }\mu\text{l}$) was applied to a normal-phase silica TLC plate, and developed 4 cm in heptane:diethyl ether:acetic acid (94:5:1) and air-dried. The plate was re-developed 8 cm in petroleum ether:acetic acid (98:2). We visualized ¹⁴C-lipids and quantitated the relative amounts of different lipid species using the Bio-Rad GS-363 phosphorimaging system.

Analysis of PA-824 metabolism

[¹⁴C]PA-824 was synthesized by condensation of [¹⁴C]-*p*-trifluoromethoxy benzyl bromide (NEN) with 3S-hydroxy-6-nitro-2H-3,4-dihydro-[2-1b]imidazopyran (USA patent

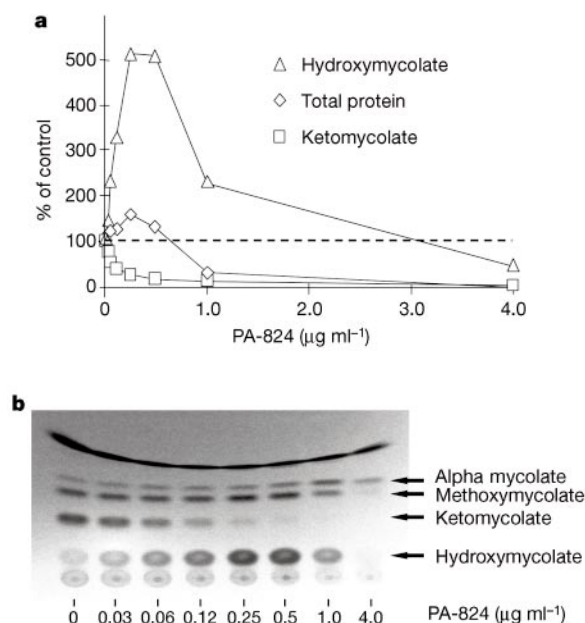


Figure 4 Effects of NAP on protein and lipid synthesis. **a**, Protein and lipid biosynthesis was analysed after a 3-h treatment of log-phase BCG cells with PA-824 at the indicated concentrations. Protein synthesis was measured by quantitating the incorporation of ³⁵S, and is expressed as a percentage of incorporation in untreated controls. **b**, Effects on lipid synthesis were analysed by incorporation of ¹⁴C-acetate and thin layer chromatography of extracted labelled lipids. Relative amounts of lipid species were determined by phosphorimager analysis and are expressed as a percentage of incorporation in untreated controls. Mycolate species were initially identified by comigration with previously purified and characterized material and further confirmed by both 500-MHz ¹H-NMR and electrospray mass spectroscopy of purified material from PA-824-treated MTB strain H37Rv12.

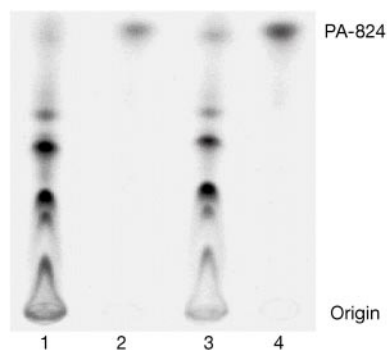


Figure 5 Metabolism of ¹⁴C-labelled PA-824 BCG cells were incubated with labelled drug, and labelled metabolites were recovered from disrupted cells and analysed by TLC. Lane 1, drug-susceptible BCG parent strain 1102; lane 2, drug-resistant mutant strain SM3; lane 3, strain SM3 transformed with plasmid pMV206-fgd; lane 4, strain SM3 transformed with control plasmid pMV206.

5668127). We treated 5 ml log-phase rolling cultures of BCG (absorbance₅₄₀ = 0.2) in 7H9 broth plus ADC enrichment cells with 25 µCi of ¹⁴C-PA-824 at a concentration of 1 µg ml⁻¹. We collected cells and prepared the lysate as above. The lysate was centrifuged for 10 min at 10,000g and the resulting supernatant evaporated to dryness. We extracted the residue with 10 µl H₂O, and applied 3 µl applied to a normal-phase silica TLC plate. The plate was developed in 1-butanol:H₂O:acetic acid (4:1:1). Radioactivity was recorded by phosphorimager.

Compound evaluation in animal infection models

NAP compounds tested in animals were formulated in 10% hydroxypropyl-β-cyclodextrin and 10% lecithins in water at a final concentration of 5 mg ml⁻¹. In mice, all drugs were administered by oral gavage. Balb-C mice were infected by tail vein injection with 10⁶ CFU of MTB-*lux* consisting of strain H37RV containing plasmid vector pMH102, a hygromycin resistant version of pMH30 (ref.11). Mycobacterial burden in mouse organ homogenates was measured by luminometry using relative light units (RLU) as a surrogate for CFUs as described¹¹. Drugs were tested in MTB infected guinea pigs essentially as described²⁴. Outbred Hartley-strain guinea pigs (Charles River Breeding Laboratories) were infected with 5–10 viable organisms per animal by aerosol with a single-cell suspension of *M. tuberculosis* H37Rv (ATCC 27294). Each animal was randomly assigned a treatment group and sacrifice interval. Drug treatment was initiated 4 weeks post-challenge. PA-824 formulated as described was mixed with an equal volume of commercial pina colada mix (Mr and Mrs T) to increase palatability, and administered at 40 mg kg⁻¹ per oral (p.o.) once daily. Each animal in the control group received a single dose of 0.5 ml INH (Isoniazid; Sigma) daily p.o. containing 25 mg kg⁻¹ INH in 50% sucrose (Sigma). Untreated animals received a single dose of 0.5 ml 50% sucrose daily p.o. as a placebo.

Guinea pigs were killed by the intraperitoneal injection of 1–3 ml of sodium pentobarbital (Sleepaway, Fort Dodge Laboratories) at 8 weeks following respiratory challenge. The thoracic cavities were opened aseptically and the right lower lobe of the lung was removed for bacterial culture. Lung tissue was homogenized in sterile Teflon-glass homogenizers in 4.5 ml of sterile physiological saline. The number of viable MTB organisms in each organ was determined by inoculating appropriate dilutions onto duplicate Middlebrook 7H10 agar plates (Remel, Lenexa, KS). The colonies were counted after 3 weeks incubation at 37 °C. Data were expressed as mean log₁₀ number of viable organisms per lung lobe. All animal infections and subsequent drug treatments were performed in a BL3 biohazard facility designed for use with class 3 human pathogens. We used analysis of variance (general linear model) to test the effects of drug treatment on tissue bacterial load. When significant treatment effects were indicated, differences between means were assessed by Duncan's multiple range test. A 95% confidence level was set for all tests.

Received 23 February; accepted 10 May 2000.

- Dye, C., Scheele, S., Dolin, P., Pathania, V. & Ravigione, M. C. Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project. *J. Am. Med. Ass.* **282**, 677–686 (1999).
- Bloom, B. R. & Small, P. M. The evolving relation between humans and *Mycobacterium tuberculosis*. *N. Engl. J. Med.* **338**, 677–678 (1998).
- Kaufman, S. & van Embden, J. Tuberculosis: a neglected disease strikes back. *Trends Microbiol.* **1**, 2–5 (1993).
- Bloom, B. R. & Murray, C. J. Tuberculosis: commentary on a reemergent killer. *Science* **257**, 1055–1064 (1992).
- Fischl, M. A. *et al.* Clinical presentation and outcome of patients with HIV infection and tuberculosis caused by multiple-drug-resistant bacilli. *Ann. Int. Med.* **117**, 184–190 (1992).
- Agrawal, K. C. *et al.* Potential radiosensitizing agents. Dinitroimidazoles. *J. Med. Chem.* **22**, 593–586 (1979).
- Nagarajan, K., Shankar, R. G., Rajappa, S., Shenoy, S. J. & Costa-Pereira, R. Nitroimidazoles XXI 2,3-dihydro-6-nitroimidazo [2,1-b] oxazoles with antitubercular activity. *Eur. J. Med. Chem.* **24**, 631–633 (1989).
- Walsh, J. S. *et al.* Structural alterations that differentially affect the mutagenic and antitrichomonal activities of 5-nitroimidazoles. *J. Med. Chem.* **30**, 150–156 (1987).
- Ashtekar, D. R. *et al.* *In vitro* and *in vivo* activities of the nitroimidazole CGI 17341 against *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **37**, 183–186 (1993).
- Wayne, L. G. & Sramek, H. A. Metronidazole is bactericidal to dormant cells of *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **38**, 2054–2054 (1994).
- Hickey, M. J. *et al.* Luciferase *in vivo* expression technology: use of recombinant mycobacterial reporter strains to evaluate antimycobacterial activity in mice. *Antimicrob. Agents Chemother.* **40**, 400–407 (1996).
- Yuan, Y. & Barry, C. E. A common mechanism for the biosynthesis of methoxy and cyclopropyl mycolic acids in *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. USA* **93**, 12828–12833 (1996).
- Samuelson, J. Why metronidazole is active against both bacteria and parasites. *Antimicrob. Agents Chemother.* **43**, 1533–1541 (1999).
- Edwards, D. I. Mechanism of antimicrobial action of metronidazole. *J. Antibiot. Chemother.* **5**, 499–502 (1979).
- Lisitsyn, N. & Lisitsyn, N. W. M. Cloning the differences between two complex genomes. *Science* **259**, 946–951 (1993).
- Cole, S. T. *et al.* Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537–544 (1998).
- Purwantini, E. & Daniels, L. Purification of a novel coenzyme F420-dependent glucose-6-phosphate dehydrogenase from *Mycobacterium smegmatis*. *J. Bacteriol.* **178**, 2861–2866 (1996).
- Purwantini, E. & Daniels, L. Molecular analysis of the gene encoding F420-dependent glucose-6-phosphate dehydrogenase from *Mycobacterium smegmatis*. *J. Bacteriol.* **180**, 2212–2219 (1998).
- Barry, C. E. *et al.* Mycolic acids: structure, biosynthesis and physiological functions. *Prog. Lipid Res.* **37**, 143–179 (1998).
- Yuan, Y., Zhu, Y., Crane, D. D. & Barry, C. E. The effect of oxygenated mycolic acid composition on cell

wall function and macrophage growth in *Mycobacterium tuberculosis*. *Mol. Microbiol.* **29**, 1449–1458 (1998).

- Holt, J. G., Krieg, N. R., Sneath, P. H. A., Staley, J. T. & Williams, S. T. in *Bergey's Manual of Determinative Bacteriology* 9th edn 597 (Williams and Wilkins, Baltimore, 1994).
- Araim, T. M., Resconi, A. E., Hickey, M. J. & Stover, C. K. Bioluminescence screening *in vitro* (Bio-Siv) assays for high-volume antimycobacterial drug discovery. *Antimicrob. Agents Chemother.* **40**, 1536–1541 (1996).
- van Embden, J. D. *et al.* Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *J. Clin. Microbiol.* **31**, 406–409 (1993).
- Smith, D. W., Balasubramanian, V. & Wiegshaus, E. H. A guinea pig model of experimental airborne tuberculosis for the evaluation of the response to chemotherapy: the effect on bacilli in the initial phase of treatment. *Tubercle* **72**, 223–231 (1991).

Correspondence and requests for materials should be addressed to C.K.S. (e-mail: kstover@pathogenesis.com).

Embryonic lethality in mice homozygous for a processing-deficient allele of Notch1

Stacey S. Huppert*, Anh Le, Eric H. Schroeter*, Jeffrey S. Mumm*, Meera T. Saxena*, Laurie A. Milner† & Raphael Kopan*

* Department of Molecular Biology and Pharmacology and the Department of Medicine (Division of Dermatology), Washington University School of Medicine, St Louis, Missouri 63110, USA

† The Fred Hutchinson Cancer Research Center and the University of Washington School of Medicine, Seattle, Washington 98109, USA

The Notch genes encode single-pass transmembrane receptors that transduce the extracellular signals responsible for cell fate determination during several steps of metazoan development. The mechanism by which extracellular signals affect gene transcription and ultimately cell fate decisions is beginning to emerge for the Notch signalling pathway. One paradigm is that ligand binding to Notch triggers a Presenilin1-dependent proteolytic release of the Notch intracellular domain from the membrane¹, resulting in low amounts of Notch intracellular domain which form a nuclear complex with CBF1/Su(H)/Lag1 to activate transcription of downstream targets². Not all observations clearly support this processing model, and the most rigorous test of it is to block processing *in vivo* and then determine the ability of unprocessed Notch to signal. Here we report that the phenotypes associated with a single point mutation at the intramembranous processing site of Notch1, Val1,744→Gly, resemble the null Notch1 phenotype^{3,4}. Our results show that efficient intramembranous processing of Notch1 is indispensable for embryonic viability and proper early embryonic development *in vivo*.

The plausibility of the processing model has been demonstrated by biochemical experiments *in vitro*^{5–7} and genetic analyses in *Drosophila*^{5,8,9}, which show that the Notch intracellular domain (NICD) gains access to the nucleus in a ligand-dependent manner. In addition, Presenilin-null mice and nematodes phenocopy animals lacking CBF1/Su(H)/Lag1 (CSL)^{10–13}, suggesting that the loss of nuclear Notch as a result of abolished Presenilin-dependent processing is the root cause of the observed phenotypes. But not all data support this model. In some transgenic flies, membrane-tethered Notch has a greater efficacy than NICD¹⁴; in addition, truncated versions of tethered Notch can signal in Presenilin-deficient flies¹⁵ and cells¹⁶. Consequently, either processing of the Notch protein is not obligate for its function or NICD is generated by other means in artificial situations (for example, truncated Notch proteins can be translated from a methionine within their transmembrane domain¹⁷, generating a NICD-like fragment independent of processing).

To address this issue, we chose a reverse genetic approach, mutating the codon GTG (valine) to GGG (glycine) at amino-acid position 1,744 in the mouse *Notch1* gene. This position is the site for proteolytic cleavage and is critical for Notch1 intracellular processing in tissue-culture cells (Fig. 1a)⁶. In activated Notch1 proteins, processing is not completely abolished in any of the single amino-acid substitutions that we tested at this time, and activity is observed when these mutant proteins are overexpressed⁶ (Fig. 1a;

E.S. and R.K., manuscript in preparation). However, ligand-dependent Notch processing is greatly reduced by these point mutations⁶. Therefore, *in vivo* Notch1 should be minimally and 'inefficiently' processed and should direct a hypomorphic phenotype if processing at this site is critical for Notch signalling activity as it is for activated forms of Notch1 protein in tissue culture⁶. No phenotype is expected if intramembranous proteolysis of Notch1 is not required for signalling.

Using the strategy outlined in Fig. 1, we generated heterozygous animals carrying two germline mutations: the V1744G mutation in exon F, and a 42-base pair (bp) pLox site insertion in the intron separating exons E and F (this allele is designated *N1*^{V→Glox}; Fig. 1b). We isolated homozygous embryos from heterozygous crosses of *N1*^{V→Glox}/+ mice at gestational day 8.5 (E8.5) until E10.5. The percentage of homozygous embryos ranged from 20 to 30%, within the expected mendelian frequency (25%; Table 1). As reported for two null alleles of *Notch1* (*N1*^{Δ1} and *N1*ⁱⁿ³²)^{3,4}, embryo absorption was detected between E10 and E12, and no homozygous embryos were recovered past E12. In a total of 12 litters, no homozygous mice were born from heterozygous crosses.

The striking similarity of embryos expressing the processing-deficient Notch1 allele to the Notch1 null embryos is consistent with the hypothesis that efficient Notch processing is necessary for the early embryonic developmental aspects of Notch activity. As homozygous *N1*^{V→Glox} embryos die, these results are consistent with the hypothesis that efficient processing of Notch 1 is essential for survival past E10 and suggest that *N1*^{V→Glox} is a severe hypomorphic allele. Alternatively, our strategy may have resulted in loss of Notch1 protein expression as a consequence of inserting foreign sequences (the 42-bp pLox site) into an intron. We prepared RNA from

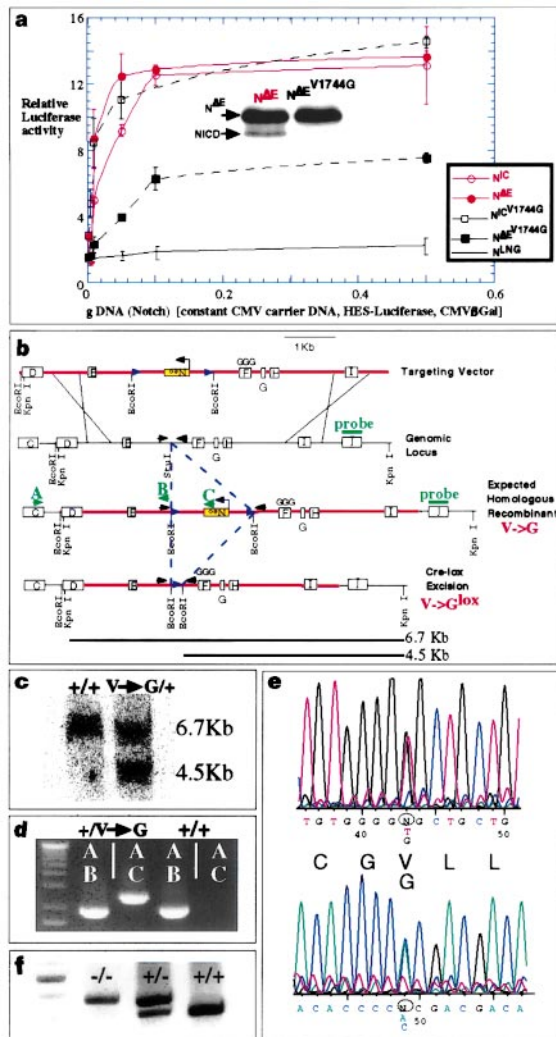


Figure 1 Strategy for generating a processing-deficient *Notch1* allele. **a**, Effects of membrane tether on Notch and the Notch processing mutant V1744G. Western blot analysis shows a reduction in the amount of NICD produced in cells expressing the *N*^{ΔE}_{V1744G} variant (at 1 μg ml⁻¹) compared with wild-type *N*^{ΔE} (inset, middle); however, the V1744G mutation does not affect signalling in a non-membrane-tethered form of Notch (*N*^{ΔE}_{V1744G}), as illustrated by HES–Luciferase reporter activity. **b**, The V1744G mutation was inserted into exon F (nomenclature from ref. 28). A floxed *neo*-selection cassette was inserted into the intron separating exons E and F in the mutated genomic fragment, and the resulting construct was used to target the Notch1 processing mutation V1744G into the *Notch1* gene. **c**, All G418-resistant ES cell clones were analysed by Southern blot to determine whether the integration was homologous. Genomic DNA was cut with *EcoRI/KpnI* and probed with a region of exon J. The expected fragments and the probe are as indicated in **b**. **d**, Homologous recombination was also verified by DNA amplification with a primer exterior to the targeting vector (A) and a *neo* primer (C). Primers A, B and C are indicated in **b**. **e**, The presence of the processing mutation was verified by sequencing genomic ES cell DNA. The *N1*^{V→Glox}/+ animals were genotyped by PCR of the intron containing the loxP site insertion. **f**, A PCR reaction distinguishing all three genotypes resulting from a cross of *N1*^{V→Glox}/+ animals is shown.

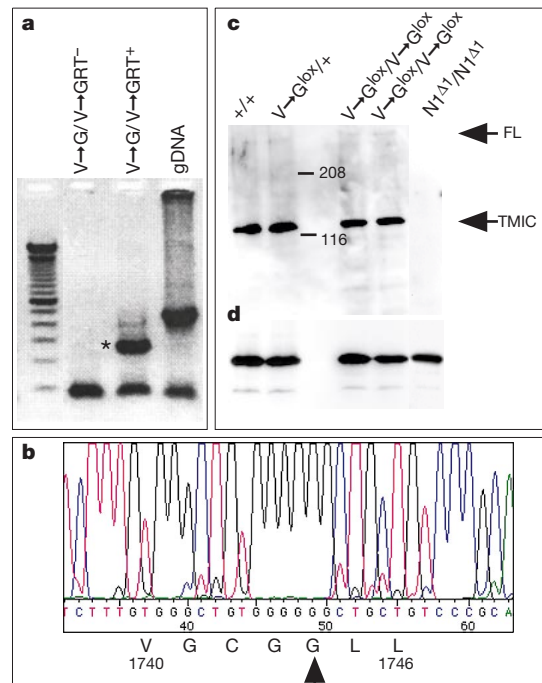


Figure 2 The targeted *Notch1* allele containing the V1744G mutation is both transcribed and translated. **a**, Example of RT–PCR of E10.5 *N1*^{V→G}/*N1*^{V→G} animals. Lane 2, RNA without RT. Lane 3, RNA with RT, producing a cDNA product of 240 bp and a gDNA product of 440 bp. **b**, Sequence of a homozygous *N1*^{V→G} embryo from the product shown by asterisk in **a**. **c**, Western blot of protein extracted from E9.5 animals (genotypes indicated) detecting Notch1 protein with anti-mN1A. Arrows indicate the position of full-length (FL) and transmembrane-tethered carboxy-terminal product of the furin cleavage containing the intracellular domain (TMIC) of Notch1 protein. **d**, The same western blot stripped and re-analysed with anti-β-catenin showing protein in all lanes.

homozygous $N1^{V-G}/N1^{V-G}$ embryos at E10.5, and carried out reverse transcription using primers within exon F and exon G, to distinguish between amplification of genomic DNA and complementary DNA generated from messenger RNA or heteronuclear RNA (hnRNA) (Fig. 2a). Sequencing the amplified messenger RNA product (Fig. 2a, asterisk) confirmed that embryos homozygous for the Notch1 intracellular processing mutation produced mRNA coding for glycine at position 1,744 (glycine, GGG, instead of valine, GTG; Fig. 2b). If most of the embryonic mRNA lacks exon F owing to aberrant splicing at the E/F junction, we would expect little if any Notch1 protein to be translated owing to frameshifts. To demonstrate that Notch1 protein is efficiently translated in $N1^{V-Glox}/N1^{V-Glox}$ embryos, equivalent amounts of total protein (as determined by β -catenin immunodetection, Fig. 2d) were analysed by western blot using an antibody raised against the CDC-NCR region of Notch1 (Fig. 2c). Comparable amounts of mature Notch1 protein are present in both the processing-deficient embryos and their heterozygous and wild-type littermates. Notch1 protein is not detected in $N1^{\Delta1}$ homozygous embryos, demonstrating the specificity of this antibody to Notch1, as homozygous null embryos do express *Notch2* (ref. 4). Thus, we conclude that the targeted Notch1 intracellular processing-deficient allele is both transcribed and translated with similar efficiency to the wild-type allele. Production of a processing-deficient Notch1 mutant protein even at wild-type levels is therefore functionally analogous to a Notch1 protein null in regard to viability.

A comparison between homozygous $N1^{V-Glox}$ and $N1^{\Delta1}$ null embryos³ might identify developmental decisions either that are independent of Notch1 processing or that can occur with inefficient Notch1 processing. We compared the phenotypes of $N1^{\Delta1}$ and $N1^{V-Glox}$ at E10.5 to determine whether embryos lacking Notch1 protein display developmental abnormalities not observed in the processing-deficient embryos (Fig. 3a–c). Morphological comparison revealed that $N1^{\Delta1}$ null and $N1^{V-Glox}$ embryos are almost indistinguishable. The development of Notch1 mutant embryos is retarded as compared with their heterozygous and wild-type littermates. $N1^{V-Glox}$ embryos at E9.5 form 19–23 somites compared with 22–26 for their heterozygous and wild-type littermates. The $N1^{\Delta1}$ embryos are even more delayed, forming only 14–18 somites by E9.5. Another prominent phenotype is a distended pericardial sac in both $N1^{\Delta1}$ and $N1^{V-Glox}$ embryos, but both develop a beating heart. Embryos of both genotypes form morphologically normal forelimb buds, and otic and optic vesicles.

A new phenotype, which may contribute to embryonic lethality, was described for Notch1 null embryos in the C57Bl6/J congenic background (ref. 18; and T. Gridley, personal communication). These embryos have severe defects in angiogenic vascular remodeling in the yolk sac at E9.5 (ref. 18). A similar yolk-sac defect is observed in embryos lacking all Presenilin activity¹¹. In all $N1^{V-Glox}$

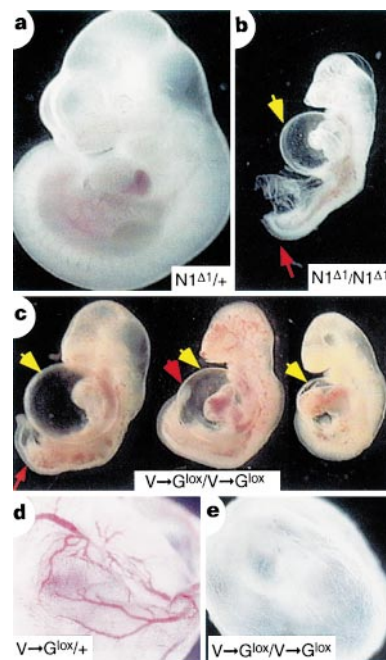


Figure 3 Overall morphological phenotypes of the following genotypes: **a**, $N1^{\Delta1}/+$; **b**, $N1^{\Delta1}/N1^{\Delta1}$ (Notch1 null); **c, e**, $N1^{V-Glox}/N1^{V-Glox}$; **d**, $N1^{V-Glox}/+$. Extra-embryonic tissue was used to genotype each embryo. Arrows point to the variable deficit in posterior development (**b** and **c**) and the expanded pericardial sac at E10.5. $N1^{V-Glox}/N1^{V-Glox}$ embryos show a lack of vascular morphogenesis¹⁸ at E9.5 (**e**) compared with their heterozygous litter mate (**d**).

embryos that we examined, the yolk sac displays similar defects to those described for Notch1 null ($N1^{in32}$)¹⁸ and the Presenilin double mutant¹¹ embryos (Fig. 3e).

We performed whole-mount *in situ* mRNA hybridization to ascertain whether there are any other detectable phenotypic differences between $N1^{V-Glox}$ and $N1^{\Delta1}$ embryos. The myotome and the sclerotome lineages (as detected by *myogenin*, Fig. 4c and data not shown; and *Uncx4.1*, Fig. 4a, b, d–g) are formed in the $N1^{V-Glox}$ as in the $N1^{\Delta1}$ embryos³. Somitogenesis is not completely disrupted in $N1^{\Delta1}$, but rather the consistent metamerically repeated pattern is poorly maintained. *Uncx4.1*, whose expression reflects this metamerically repeated pattern in the caudal compartment of the sclerotome¹⁹, further illustrates this point (Fig. 4a, b, d–g). Disruptions in size and bilateral symmetry of the somites are observed in $N1^{\Delta1}$ null embryos (Fig. 4a, b). The anterior somite phenotype of the $N1^{V-Glox}$ embryos shows similar disruptions in the metamerically repeated pattern (Fig. 4c), but at lower penetrance than in the $N1^{\Delta1}$. No irregularities in somite size were detected in the posterior region of the 22 $N1^{V-Glox}$ embryos that we analysed (Fig. 4e, f). Neural tube kinks are observed in the $N1^{\Delta1}$ embryos, extending posterior to the forelimb (Fig. 4a). Whereas only 1 out of 23 wild-type embryos displayed similar kinks in the anterior region, 14 out of 22 $N1^{V-Glox}$ embryos displayed neural tube kinks anterior to the forelimb (Fig. 4c, g). No neural tube defects are seen posterior to the forelimb (Fig. 4e, f). Thus, in the anterior region of both mutant Notch1 alleles, the phenotype is comparable albeit at lower penetrance.

Alteration in ligand expression is detected in $N1^{\Delta1}$ embryos. *Delta1* is expressed in the presomitic mesoderm, the caudal half of each somite, the forebrain and the neuroepithelium of the presumptive midbrain region, and in isolated cells in the neural tube²⁰. In E8.5 embryos, *Delta1* expression in all *Notch1* alleles is almost indistinguishable from wild type (Fig. 5a–c); however, as shown previously²¹, *Delta1* expression is upregulated in the neural tube in E8.5 $N1^{\Delta1}$ embryos (Fig. 5c), although to a lesser extent than

Table 1 The number of homozygous $N1^{V-Glox}$ embryos at embryonic day 8.5 to birth

$V-Glox/+ \times V-Glox/+$	$+/+$	$V-Glox/+$	$V-Glox/V-Glox$
+			
E8.5	7 (28%)	13 (52%)	5 (20%)
Expected	6 (25%)	13 (50%)	6 (25%)
E9.5	40 (24%)	77 (46%)	50 (30%)
Expected	42 (25%)	83 (50%)	42 (25%)
10.5	22 (29%)	38 (50%)	16 (21%)
Expected	19 (25%)	38 (50%)	19 (25%)
E12.5	5 (36%)	5 (36%)	4 (28%)*
Expected	3.5 (25%)	7 (50%)	3.5 (25%)
At birth	37 (37%)	64 (63%)	0
Expected	25 (25%)	51 (50%)	25 (25%)

The percentage of each genotype is shown with the expected numbers of each genotypic class given in parentheses. E designates embryonic day.

* Absorbed embryos.

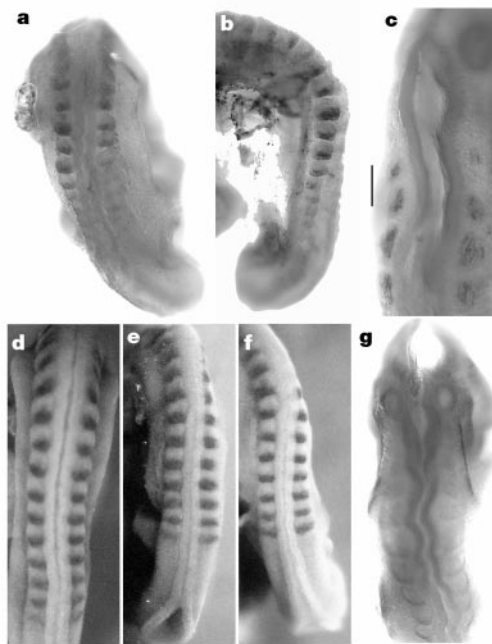


Figure 4 *In situ* hybridization analysis of E9.5 animals. Genotypes: **a, b**, $N1^{\Delta 1}/N1^{\Delta 1}$; **c, e–g**, $N1^{V \rightarrow Glax}/N1^{V \rightarrow Glax}$; **d**, wild type. All panels show hybridizations with an *Uncx4.1* RNA probe was used except in **c** where a *myogenin* RNA probe was used. Compare the posterior view of somitogenesis (**a, b, d–f**). A disturbance in patterning of anterior somites (indicated by the lines) is shown in **c**. Neural tube kinks in the anterior of $N1^{V \rightarrow Glax}/N1^{V \rightarrow Glax}$ embryos are shown in **c** and **g**; no neural tube kinks are detected in the posterior of $N1^{V \rightarrow Glax}/N1^{V \rightarrow Glax}$ embryos (**e, f**), but kinks are visible in the posterior region of $N1^{\Delta 1}/N1^{\Delta 1}$ embryos (**a**).

in CSL^{RBPjk} null embryos or Presenilin1/Presenilin2 double homozygotes¹⁰. We find similar *Delta1* upregulation in E9.5 $N1^{V \rightarrow Glax}$ embryos in anterior regions (Fig. 5f, inset). As in somitogenesis, the posterior effect is less pronounced (Fig. 5e, f). Therefore, a processing-defective Notch1 receptor has a similar effect on *Delta1* expression in the anterior region as observed in a Notch1 null background²¹. We conclude that the overall phenotypes of $N1^{V \rightarrow Glax}$ and $N1^{\Delta 1}$ null are qualitatively equivalent, with some quantitative differences depending on the region analysed.

Why is there a difference between anterior and posterior regions in the $N1^{V \rightarrow Glax}$ embryos? In the posterior region, where *Notch1* expression is the highest in the embryo²², inefficient proteolysis may still produce NICD above threshold levels (Fig. 1a)⁶, resulting in restoration of normal morphology. We attempted to determine to what extent Notch1 is processed in the posterior of homozygous $N1^{V \rightarrow Glax}$ embryos. Because our V1744G point mutation results in a ninefold reduction in Notch1 processing (Fig. 1a), we reasoned that 10 wild-type embryos will represent 100 mutant embryos, the upper limit of what we can produce and process. Although we have been able to enrich for NICD efficiently by immunoprecipitation with a CSL antibody (provided by R. Aguilera) from 293T cells transfected with a membrane-tethered, truncated, active Notch1 ($N^{\Delta E}$)¹⁷, we have been unsuccessful in using the same strategy to purify NICD from pools of 10 wild-type E9.5 embryos. This result may reflect the small number of cells in which Notch has been activated at any given time, the short half-life of NICD *in vivo*²³, or a combination of both. This finding is also consistent with the observation that very low nuclear concentrations of Notch1 are sufficient for productive signalling⁶. Our failure to detect NICD in wild-type embryos therefore leaves the hypothesis offered above as the most likely, but unproved, explanation for improved posterior development in $N1^{V \rightarrow Glax}$ embryos.

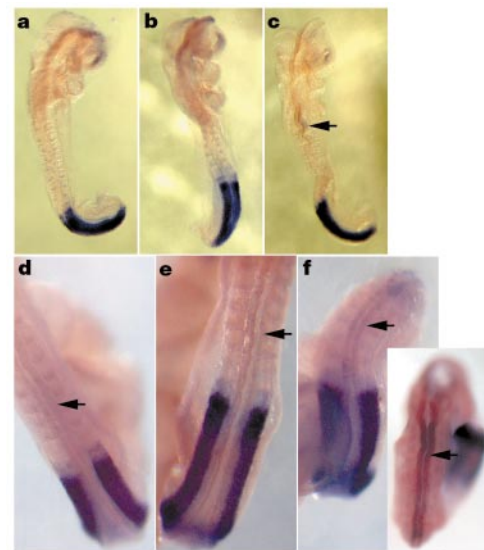


Figure 5 *In situ* hybridization of *Delta1* (*Dl1*). **a–c**, E8.5; **d–f**, E9.5. The embryos are wild type (**a, d**), $N1^{V \rightarrow Glax}/N1^{V \rightarrow Glax}$ (**b, e, f**) and $N1^{\Delta 1}/N1^{\Delta 1}$ (**c**). Arrows point to *Dl1* expression in the neural tube. Note the difference in the level of *Dl1* expression between the anterior and posterior views of the $N1^{V \rightarrow Glax}/N1^{V \rightarrow Glax}$ embryo (**f**).

In conclusion, our data show unequivocally that efficient Notch1 intracellular processing is required for viability and proper formation of yolk-sac vasculature in the mouse embryo. A processing-independent Notch1 signalling event would be predicted to function in a manner indistinguishable from that of wild-type Notch. Our analysis of mice expressing the processing-deficient Notch1 allele revealed that all the phenotypes described in the null genotypes were qualitatively reproduced. However, quantitative differences were described: the null-like phenotype in the anterior occurs at lower penetrance. The quantitative differences that we observe in the regions of high *Notch1* expression may reflect either residual processing observed in V1744G mutants, or proteolysis-independent Notch1 signalling acting at the posterior but not the anterior of the embryo. The latter possibility is unlikely given that no fundamental differences between anterior and posterior somitogenesis are known to exist. Collectively, our data establish the processing-deficient allele as a strong hypomorph of Notch1. It remains to be determined whether late functions of Notch1 also depend on efficient cleavage at V1744. □

Methods

Gene targeting Notch1 intracellular processing mutation

Genomic clones containing exon D up to exon J of *Notch1* were isolated from a library made by R.K. The library was constructed as a *Sau3AI* partial, cloned into lambda DASH (Stratagene) using *Bam*HI. The library was probed with the ankyrin region of Notch1. A clone of roughly 10 kilobases was subcloned into BlueScript. We generated a subclone spanning the region from exon D to J into which we inserted a *neo*-selection cassette disrupting the *StuI* site in the intron between exon E and F. The processing mutation was cloned into this vector in two steps (details available upon request). The vector was linearized with *SalI* and electroporated into RW4 embryonic stem cells (T. Ley; <http://medicine.wustl.edu/~escor/>). Correctly targeted clones were verified by Southern blot analysis using an external probe from exon J and by polymerase chain reaction (PCR) with a primer located 5' outside of the targeting vector ([A]exon C, 5'-CGG CCG GCT GCC TCT TTG-3', and [C]Neo, 5'-TCG CCT TCT ATC GCC TTC TTG-3'). Embryonic stem (ES) cell clones 44 and 93 were injected into C57Bl6/J blastocysts (Jackson Laboratory) to obtain germline transmission.

Heterozygous animals transmitting the targeted allele, as verified through PCR, were sequenced to confirm that the targeted allele contained the processing mutation using primers Motch 33, 5'-CAG AGA TCT CTG CAC CTC ATG TAC GTG-3', and Motch 49, 5'-CGA ACT AGT CTC TGA CAC TTT GAA ACC-3'. Embryos containing the *neo*-selection cassette from both ES cell clones (44 and 93) were initially analysed and displayed identical phenotypes. In these animals, the *neo*-expression cassette inserted in reverse orientation to the Notch transcript encodes a cryptic splice acceptor site, which can

strongly reduce the amount of RNA produced from the targeted allele²⁴. To confine the analysis to the effects of reduced Notch1 processing, we mated $N1^{V-G/+}$ mice with mice containing *cre*-recombinase under the control of the β -actin promoter²⁵ (provided by G. Martin), generating a second line of mice, in an out-bred background, in which the *neo*-selection cassette was excised ($V \rightarrow G^{lox}$ chromosome; derived from ES cell line 93). These animals were identified by the 42-bp *EcoRI* footprint that remained. This allele is designated as Notch1^{Im1Kop}.

Transfection, reporter gene assays and co-precipitations

These were done as described⁶.

Analysis of mutant embryos

Embryos were genotyped by PCR of yolk-sac DNA using the following primer sets: EintF5', 5'-TGG CGG GTC ACA GGC AGG CGG TCT C-3', and Neo, 5'-TCG CCT TCT ATC GCC TTC TTG-3', which amplifies a 412-bp fragment in the processing mutant allele ($N1^{V-G}$); EintF5' and EintF3', 5'-AAG GAT TTA GGG GGA TTT TCA AGA TAC AA-3', which amplifies a 340-bp fragment wild-type allele, and a 382-bp fragment in the *neo*-excised processing mutant allele ($N1^{V-Glox}$). *In situ* mRNA hybridization was performed as described²⁶. The probes used for hybridization were *Uncx4.1* (ref. 19), *myogenin* (provided by M. J. Thayer) and *Delta1* (provided by D. Henrique). Whole-mount immunohistochemistry was done as described³.

RT-PCR

RNA was isolated from 10.5 day $N1^{V-G}$ embryos using TRIzol Reagent (Gibco BRL). Reverse transcription was carried out using an Enhanced Avian RT-PCR kit (Sigma) and primers within exon F (Moth33) and exon G (Moth94, 5'-CCA CTC GTT CTG ATT GTC GTC CAT-3'). The cDNA amplifies a fragment of 240 bp and the hnRNA and genomic DNA amplifies a fragment of 440 bp. The 240-bp fragment was gel purified and sequenced using Moth33 and Big Dye cycle sequencing (Nucleic Acid Chemistry).

Western analysis

Single embryos were homogenized in 35 μ l of Laemmli buffer with 0.01 M dithiothreitol. The extract was heated to 65 °C for 30 min and then stored at -80 °C. The extracts were run on a 6% SDS-PAGE gel and quantified by Coomassie blue. Western analysis was done as described⁶. Notch1 protein was detected by mN1 antibody. β -catenin was detected by an antibody produced by Transduction Laboratories.

mN1A monoclonal antibody production

The mN1A mouse IgG1 monoclonal antibody, generated against the cdc10-NCR region of mNotch1 (mN1-ICADOP), was produced using standard methods of immunization and hybridoma production. mN1-ICADOP cDNA was cloned into pGEX-5X (Promega), expressed as a GST fusion protein, and affinity-purified protein was used²⁷ to immunize mice. Splenocytes from animals producing mN1-specific antibodies were fused to SP2/0 cells and resulting hybridomas screened by ELISA for anti-mN1 secretion. Individual hybridoma clones were injected IP into syngeneic mice, and purified antibody was obtained by Protein A-Sepharose affinity chromatography of the ascites fluids produced. A detailed description will be published elsewhere (L.M. *et al.* in preparation).

Received 7 January; accepted 11 May 2000.

- Chan, Y. M. & Jan, Y. N. Roles for proteolysis and trafficking in Notch maturation and signal transduction. *Cell* **94**, 423–426 (1998).
- Jarriault, S. *et al.* Signalling downstream of activated mammalian Notch. *Nature* **377**, 355–358 (1995).
- Conlon, R. A., Reaume, A. G. & Rossant, J. *Notch1* is required for the coordinate segmentation of somites. *Development* **121**, 1533–1545 (1995).
- Swiatek, P. J., Lindsell, C. E., Amo, F. F. d., Weinmaster, G. & Gridley, T. *Notch1* is essential for postimplantation development in mice. *Genes Dev.* **8**, 707–719 (1994).
- Kidd, S., Lieber, T. & Young, M. W. Ligand-induced cleavage and regulation of nuclear entry of Notch in *Drosophila melanogaster* embryos. *Genes Dev.* **12**, 3728–3740 (1998).
- Schroeter, E. H., Kisslinger, J. A. & Kopan, R. Notch-1 signalling requires ligand-induced proteolytic release of intracellular domain. *Nature* **393**, 382–386 (1998).
- De Strooper, B. *et al.* A presenilin-1-dependent gamma-secretase-like protease mediates release of Notch intracellular domain. *Nature* **398**, 518–521 (1999).
- Lecourtis, M. & Schweisguth, F. Indirect evidence for *Delta*-dependent intracellular processing of Notch in *Drosophila* embryos. *Curr. Biol.* **8**, 771–774 (1998).
- Struhl, G. & Adachi, A. Nuclear access and action of Notch in vivo. *Cell* **93**, 649–660 (1998).
- Donoviel, D. B. *et al.* Mice lacking both presenilin genes exhibit early embryonic patterning defects. *Genes Dev.* **13**, 2801–2810 (1999).
- Herreman, A. *et al.* Presenilin 2 deficiency causes a mild pulmonary phenotype and no changes in amyloid precursor protein processing but enhances the embryonic lethal phenotype of presenilin 1 deficiency. *Proc. Natl Acad. Sci. USA* **96**, 11872–11877 (1999).
- Li, X. & Greenwald, I. HOP-1, a *Caenorhabditis elegans* presenilin, appears to be functionally redundant with SEL-12 presenilin and to facilitate LIN-12 and GLP-1 signaling. *Proc. Natl Acad. Sci. USA* **94**, 12204–12209 (1997).
- Westlund, B., Parry, D., Clover, R., Basson, M. & Johnson, C. D. Reverse genetic analysis of *Caenorhabditis elegans* presenilins reveals redundant but unequal roles for sel-12 and hop-1 in Notch pathway signaling. *Proc. Natl Acad. Sci. USA* **96**, 2497–2502 (1999).
- Artavanis-Tsakonas, S., Rand, M. D. & Lake, R. J. Notch signaling: cell fate control and signal integration in development. *Science* **284**, 770–776 (1999).
- Ye, Y. H., Lukinova, N. & Fortini, M. E. Neurogenic phenotypes and altered Notch processing in *Drosophila* Presenilin mutants. *Nature* **398**, 525–529 (1999).

- Berechid, B. E., Thinakaran, G., Wong, P. C., Sisodia, S. S. & Nye, J. S. Lack of requirement for Presenilin1 in Notch1 signaling. *Curr. Biol.* **9**, 1493–1496 (1999).
- Kopan, R., Schroeter, E. H., Weintraub, H. & Nye, J. S. Signal transduction by activated mNotch: importance of proteolytic processing and its regulation by the extracellular domain. *Proc. Natl Acad. Sci. USA* **93**, 1683–1688 (1996).
- Krebs, L. T. *et al.* Notch signaling is essential for vascular morphogenesis in mice. *Genes Dev.* **14**, 1343–1352 (2000).
- Neidhardt, L. M., Kispert, A. & Herrmann, B. G. A mouse gene of the paired-related homeobox class expressed in the caudal somite compartment and in the developing vertebral column, kidney and nervous system. *Dev. Genes Evol.* **207**, 330–339 (1997).
- Bettenhausen, B., de Angelis, M. H., Simon, D., Guenet, J. -L. & Gossler, A. Transient and restricted expression during mouse embryogenesis of *Dill1*, a murine gene closely related to *Drosophila Delta*. *Development* **121**, 2407–2418 (1995).
- de la Pompa, J. L. *et al.* Conservation of the Notch signalling pathway in mammalian neurogenesis. *Development* **124**, 1139–1148 (1997).
- Reaume, A. G., Conlon, R. A., Zirngibl, R., Yamaguchi, T. P. & Rossant, J. Expression analysis of a *Notch* homologue in the mouse embryo. *Dev. Biol.* **154**, 377–387 (1992).
- Schweisguth, F. Dominant-negative mutation in the $\beta 2$ and $\beta 6$ proteasome subunit genes affect alternative cell fate decisions in the *Drosophila* sense organ lineage. *Proc. Natl Acad. Sci. USA* **96**, 11382–11386 (1999).
- Carmeliet, P. *et al.* Abnormal blood vessel development and lethality in embryos lacking a single VEGF allele. *Nature* **380**, 435–439 (1996).
- Lewandoski, M. & Martin, G. R. Cre-mediated chromosome loss in mice. *Nature Genet.* **17**, 223–225 (1997).
- Yuan, W. *et al.* The mouse SLIT family: Secreted ligands for ROBO expressed in patterns that suggest a role in morphogenesis and axon guidance. *Dev. Biol.* **212**, 290–306 (1999).
- Milner, L. A. *et al.* Inhibition of granulocytic differentiation by mNotch1. *Proc. Natl Acad. Sci. USA* **93**, 13014–13019 (1996).
- Girard, L. *et al.* Frequent provirus insertional mutagenesis of Notch1 in thymomas of MMTVD/myc transgenic mice suggests a collaboration of c-myc and Notch1 for oncogenesis. *Genes Dev.* **10**, 1930–1944 (1996).

Acknowledgements

We thank T. Ley, E. Ross, D. Ornitz and S. Hua for technical help; R. Conlon and T. Gridley for providing animals and communicating results before publication; R. Aguilar, D. Henrique, B. Herrmann, G. Martin, J. Sanes, M. Thayer and S. Troyanovsky for reagents; R. Cagan and O. Pourquie for reading and commenting on the manuscript; and members of the Kopan lab—X. Tian, A. Nichols, M.-H. Lin, J. Kisslinger, B. Hadland and J. Books. This work was supported by the NIH.

Correspondence and requests for materials should be addressed to R.K. (e-mail: kopan@molecool.wustl.edu).

An extracellular activator of the *Drosophila* JAK/STAT pathway is a sex-determination signal element

Louise Sefton, John R. Timmer, Yan Zhang, Florence Béranger & Thomas W. Cline

Department of Molecular and Cell Biology, 401 Barker Hall, University of California, Berkeley, California 94720-3204, USA

Metazoans use diverse and rapidly evolving mechanisms to determine sex. In *Drosophila melanogaster* an X-chromosome-counting mechanism determines the sex of an individual by regulating the master switch gene, *Sex-lethal* (*Sxl*)¹. The X-chromosome dose is communicated to *Sxl* by a set of X-linked signal elements (XSEs), which activate transcription of *Sxl* through its 'establishment' promoter, *Sxl_{pe}*. Here we describe a new XSE called *sisterlessC* (*sisC*) whose mode of action differs from that of previously characterized XSEs, all of which encode transcription factors that activate *Sxl_{pe}* directly. In contrast, *sisC* encodes a secreted ligand for the *Drosophila* Janus kinase (JAK) and 'signal transducer and activator of transcription' (STAT) signal transduction pathway and is allelic to *outstretched* (*os*, also called *unpaired*). We conclude that *sisC* works indirectly on *Sxl* through this signalling pathway because mutations in *sisC* or in the genes encoding *Drosophila* JAK or STAT reduce expression of *Sxl_{pe}*.

similarly. The involvement of *os* in sex determination confirms that secreted ligands can function in cell-autonomous processes. Unlike sex signals for other organisms, *sisC* has acquired its sex-specific function while maintaining non-sex-specific roles in development, a characteristic that it shares with all other *Drosophila* XSEs².

The two copies of XSEs present in XX individuals in *Drosophila* specify female development by transiently activating *Sxl_{pe}* in the young embryo. A positive autoregulatory feedback loop acting on RNA splicing keeps *Sxl* active in females thereafter. Male development ensues in XY individuals because their single set of XSEs is insufficient to activate *Sxl_{pe}*. Because *Sxl* controls the vital process of X-chromosome dosage compensation as well as sex determination, sexually inappropriate expression of *Sxl* is lethal. For example, simultaneous duplication of *sisA* and *sisB* kills males, as a female dose of these two XSEs in males causes *Sxl* to be expressed in its female mode, thereby reducing X-linked gene expression.

Because XSEs act additively, males that would be killed by an excess dose of one group of XSEs can be rescued by compensating mutations that reduce the dose of other XSEs. A genetic screen³ based on this principle of additivity generated five new mutations that, as we show here, define a new XSE, *sisC*. The first four *sisC* alleles recovered, including an apparent null, *sisC¹*, had no phenotype by themselves, even *in trans* to deficiencies of the region. In contrast, *sisC⁵*, which is also null for sex-determination, exhibited phenotypes unrelated to sex: variably reduced viability (females more than males), female sterility and tergite defects. All mutations were mapped by recombination and deficiency analysis close to *os* based on their interactions with mutations in other XSEs.

We show below that *sisC* and *os* are the same gene, but this possibility seemed to have been excluded by the initial characterization of the putative *sisC* region. We, like others^{3–5}, found that *Df(1)os^{1a}* seemed to be wild-type for sex-determination function. Moreover, construction of an *os⁵sisC¹* chromosome placed *sisC* centromere-proximal to *os* and the phenotype of the double mutant adults gave no hint that both lesions affected the same gene. Using DNA centromere-proximal to *os* from ref. 6, we identified a DNA breakpoint for the atypical allele *sisC⁵* precisely where the genetics had predicted *sisC* to be—outside the region deleted by *Df(1)os^{1a}*; however, we suspected complications when we found no candidate *sisC* RNAs or other *sisC* lesions near this breakpoint.

Our subsequent discovery³ that embryonic lethal *os* alleles (*os^{upd}* by convention) failed to complement *sisC⁵* female sterility and tergite defects indicated that the non-sex-determination defects of this atypical *sisC* allele must be due to a different lesion that disrupted *os* slightly. Our next finding³ that embryonic lethal *os* alleles were deficient for XSE function indicated that this second lesion in *sisC⁵* might also be responsible for the sex-determination defect. This finding coincided with the recovery of a candidate *os* complementary DNA (λ KZ-GR)⁷.

With this cDNA as a probe, we found the anticipated second change in *sisC⁵*, 100 kilobases (kb) centromere-distal to the first and within *Df(1)os^{1a}*, just upstream of the 5' end of λ KZ-GR (Fig. 1b). DNA sequencing showed *sisC⁵* to be a simple inversion, and allowed us to design polymerase chain reaction (PCR) primers that would amplify only *os^{upd}* DNA from *sisC⁵/os^{upd}* females. Our discovery that *os^{upd-3}* and *os^{upd-4}* were a 2-base-pair (bp) insertion at codon 143 and a C-to-T change causing a stop at codon 60, respectively, showed that λ KZ-GR corresponded to *os*, as confirmed by ref. 7.

Characterization of *os* DNA for the other *sisC* mutants (Fig. 1b) showed that *sisC* and *os* are the same gene and that the first *sisC⁵* breakpoint and the *Df(1)os^{1a}* XSE test results were misleading. *sisC¹* is a 1.6-kb DNA insertion and a ~100-bp duplication just upstream of λ KZ-GR. All weaker *sisC* lesions also fell within *os*: *sisC³* deletes 7 bp upstream of the translation start, *sisC⁴* inserts tyrosine after alanine 128, and *sisC²* substitutes CGG for GTT 5 bp downstream of a 5' splice site (decreasing RNA splicing efficiency as determined by reverse transcription (RT)-PCR, data not shown). We resolved the contradiction that *Df(1)os^{1a}* was *sisC⁺* by showing that the *Df* chromosome carried a closely linked, maternal-effect suppressor³. We now designate mutations that only disrupt sex determination as *os^{sisC}*, but refer to the locus as *sisC* when discussing *os* as an XSE.

We defined the 5' end of *os* transcripts to guide our construction of transgenes essential for showing the role of *os* in sex determination and for understanding sex-specific mutations in a non-sex-specific gene. 5' rapid amplification of cDNA ends (RACE) indicated two potential transcription start sites 1,316 and 1,439 bp upstream of the 5' end of λ KZ-GR, with the product for the +1,316 species terminating in a non-coding G which could correspond to the 5' methyl cap of a bona fide messenger RNA end. RNase protection (Fig. 2) showed that the two RACE products corresponded to the major and minor *os* mRNA species present during

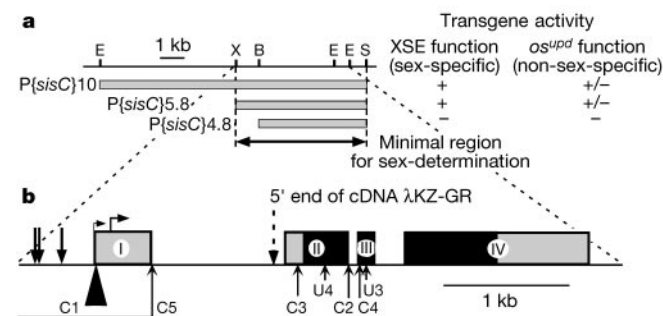


Figure 1 The *sisC/os* transcription unit. **a**, Transgenes carrying various genomic fragments (shaded boxes) were assayed for their XSE (*sisC*) activity (Table 1) and their ability to complement the non-sex-specific lethality of *os^{upd-3}* (+ full; +/- partial; - none). Restriction sites are *EcoR*1 (E), *Xba*I (X), *Bam*H1 (B) and *Sal*I (S). **b**, Roman numerals represent exons (boxes); black fill indicates open reading frame. Right arrows, two transcription start sites (Fig. 2). The larger arrow indicates the site more used. The position of the probable artefactual 5' end of the published⁷ *os* cDNA, λ KZ-GR, is shown. Down arrows, three 'XSE box' DNA sequences. Mutations in non-sex-specific lethal alleles *os^{upd-3}* and *os^{upd-4}* are labelled U3 and U4, respectively; those in the five viable sex-determination mutants (*sisC*) are labelled C1–C5. Note that the viable allele *In(1)os^{sisC-5}* eliminates *os* promoters.

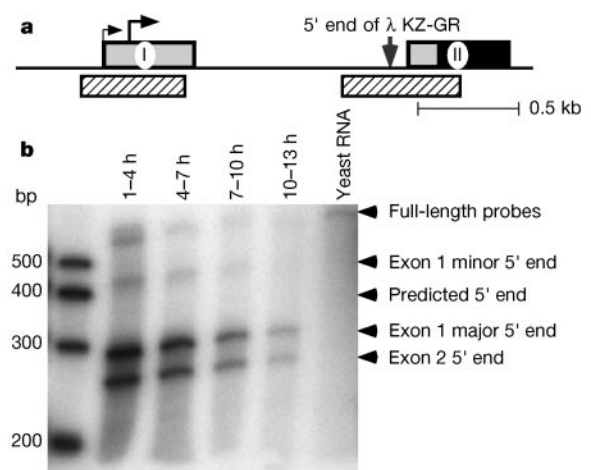


Figure 2 RNase protection analysis defining the *os* 5' end. **a**, The two probes used (stippled boxes) are indicated relative to exons I and II. **b**, A blot of the protected fragments generated from these two probes hybridized to poly(A)⁺ RNA from staged embryos of the ages indicated.

the first half of embryonic development. Sequencing the RACE products, other RT-PCR products and genomic DNA revealed the true first exon and redefined the start of the exon designated II in Fig. 1b. The 5' end of λ KZ-GR may be artefactual, as it was not found in mRNA from embryos.

Identification of these transcription start sites showed that *sisC* fits a pattern established for all other XSEs that control *Sxl* throughout the embryo: *sisC* has three copies of the sequence CAGGTAG less than 0.5 kb upstream of its first transcription start site (-242, -424 and -450 bp, Fig. 1b). Several conserved copies of this sequence or its complement were found previously upstream of the transcription start sites for *sisA*, *sisB* and *Sxl*, leading to speculation that this sequence might be involved in the unusually early onset of XSE transcription⁸.

Three genomic transgenes were constructed with different amounts of 5' sequence but the same 3' end (Fig. 1a and Table 1). Lines with 6.5 kb or 0.6 kb of genomic sequence upstream of the transcription start sites (*P{sisC}10* and *P{sisC}5.8*, respectively) provided high XSE function but only partially complemented *os*^{upd} lethals. In contrast, transgenes (*P{sisC}4.8*) lacking the transcription start sites and most of exon 1, but still containing 43 bp more uninterrupted 5' sequence than *In(1)os*^{sisC-5}, had no XSE and *os* activity. Hence, the significant level of *os*⁺ activity provided by *In(1)os*^{sisC-5} despite its disrupted transcription start sites is unlikely to be due to an undiscovered endogenous *os* promoter, but rather to

introduction of a foreign promoter providing activity sufficient only for non-sex-specific functions.

The data in Table 1 show that *os* is an XSE. To merit XSE status, lowering the zygotic dose of the gene must decrease the probability of *Sxl*⁺ activation in females, but increasing the dose of the same gene must also increase that probability in males. Table 1(a) shows that the two larger *sisC* transgenes can compensate for mutations that lower perceived X-chromosome dose in females. Females simultaneously homozygous for a *sisC* null and heterozygous for a mutation in *sisA* would die but are rescued by the two larger transgenes. Line-to-line variation in rescue probably reflects the effects of different insertion-site positions. Table 1(b) shows that *os*⁺ transgenes can kill males by increasing perceived X-chromosome dose. Males were sensitized to increased *sisC*⁺ by an increase in *sisA*⁺ dose. A double dose of *sisC*⁺ is more lethal to males than a single dose, reflecting the additive behaviour expected of XSEs. The expected *Sxl* dependence of this lethality is shown by the fact that *Sxl*⁻ males are fully viable regardless of XSE dose. The effect of *sisC*⁺ transgenes is comparable to that of a chromosomal duplication of *sisC*⁺, *Dp(1;3)JC153* (data not shown), but is smaller than seen for extra doses of *sisA*⁺ or *sisB*⁺.

As is true for other *Drosophila* XSEs, eliminating *sisC* activity reduces expression of *Sxl*_{pe}, but less than eliminating *sisA* or *sisB*. Like *sisA* and *sisB* mutations, but unlike *run* mutations, *sisC* mutations affect *Sxl*_{pe} throughout the embryo. To reach these conclusions, we exploited a *Sxl*_{pe}-*lacZ* reporter transgene that mimics the endogenous *Sxl* promoter, only expressing β -galactosidase in females^{9,10}. Expression was assayed by immunostaining young embryos with an anti- β -galactosidase antibody. The effects of *os* null alleles on *Sxl*_{pe} were revealed among progeny from the cross *Df(1)os*^{ue69}/FM7, *os*⁺; *P{Sxl_{pe}:lacZ}*3.0/*P{Sxl_{pe}:lacZ}*3.0 ♀ × ♂♂ *os*^{upd-3}/fcl⁺, *os*⁺ shown in Fig. 3j. With a *ftz*:*lacZ* marker on FM7, we distinguished *Df(1)os*^{ue69}/*os*^{upd-3} (null) embryos from their (striped) FM7/*os*^{upd-3} (*os*⁺) sisters. Only embryos without FM7 (*n* = 731) are included in Fig. 3j.

In non-mutant situations, most *Sxl*_{pe}:*lacZ* females stain darkly (Fig. 3a,b) and comprise the expected 50% of the progeny. The other 50% are males (Fig. 3c) whose light staining matches that of embryos lacking *P{Sxl_{pe}:lacZ}*. Most *Df(1)os/os*⁻ females in Fig. 3j stained lighter than *os*⁺ female controls but darker than males, showing that *os*⁻ generally reduces but does not eliminate *Sxl*_{pe} expression. The reduction was not always uniform across the embryo (Fig. 3d-f). Any region could be affected, but anterior expression seemed to be reduced the least. The range of effects was considerable: as fewer than 50% of the mutant embryos stained above background, some mutant females may not have expressed *Sxl*_{pe} at all, but *Sxl*_{pe} expression in a few others matched that of their *os*⁺ sisters. In contrast to the *Df(1)os/os*^{upd} females in Fig. 3j, *Df(1)os/os*^{sisC-1} females are fully viable, but they showed a comparable reduction in *Sxl*_{pe} expression (data not shown). This observation confirms that *os*^{sisC-1} is near null with respect to sex determination, but still supports normal development.

If *os* acts on *Sxl*_{pe} indirectly through effects on *Drosophila* JAK (encoded by *hopscotch* (*hop*)) and on *Drosophila* STAT (encoded by *Stat92E*), then the effect on *Sxl*_{pe} of eliminating either *hop* or *Stat92E* should be the same as eliminating *os*. We confirmed this prediction. Because only maternal rather than zygotic *hop* and *Stat92E* are likely to be relevant at the very early embryonic stage when *Sxl*_{pe} is activated¹¹⁻¹³, we eliminated the maternal contribution of these two genes by inducing homozygous mutant germline clones in mothers heterozygous for null alleles. Expression of *Sxl*_{pe}:*lacZ* in these experimentals was compared with that for control embryos derived from *hop*^{+/+} and *Stat92E*^{+/+} germ cells.

Loss of maternal *hop*⁺ did not eliminate *Sxl*_{pe} expression, but expression was substantially reduced: although 49% of the experimental embryos expressed *Sxl*_{pe}:*lacZ*, essentially identical to the 50% figure for the controls (Fig. 3k), 32% of the experimental

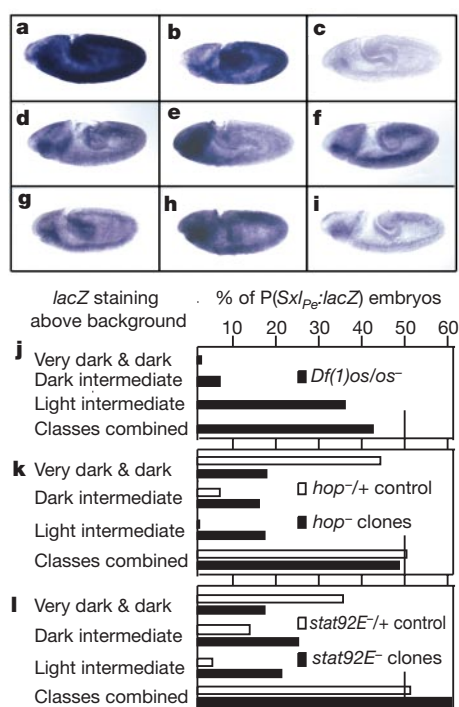


Figure 3 Mutations in the *Drosophila* JAK/STAT signalling pathway reduce expression of the *Sxl* 'establishment' promoter *Sxl*_{pe} throughout the embryo. Embryos 3–5 h after laying were immunostained for β -galactosidase derived from a *Sxl*_{pe}:*lacZ* reporter transgene that is normally on in females (a, dark; b, dark intermediate) and off in males (c, light). j, Effects on *Sxl*_{pe} expression of eliminating *os*. The line at 50% indicates expectations if all females and only females express *Sxl*_{pe}. Representative females of the 'light intermediate' class illustrate spatial variation in the effects on *Sxl*_{pe} of loss of zygotic *os*: d, effects are relatively uniform; e, posterior bias; and f, anterior bias. k, Effect of eliminating maternal *hop* (*hop*²/*hop*² germline clones) compared with a genetically related *hop*⁺ germ-cell control group. g, A 'light intermediate' female from a *hop*⁻ egg. l, Effect of loss of maternal *Stat92E* (*Stat92E*⁶³⁴⁶/*Stat92E*⁶³⁴⁶ germline clones) compared with related *Stat92E*⁺ controls. h, A 'light intermediate' female from a *Stat92E*⁺ egg. i, A control male from l showing localized *lacZ* expression from the enhancer trap in *Stat92E*⁶³⁴⁶.

(a) *sisC*⁺ transgenes rescue females that have reduced XSE activity

* Crosses were of the form $w \ v \ sisA^1 \ m \ Bx^2 \ sisC^1 \ car/Binsinscy$, $y \ w \ sn^{v2} B \ \varnothing \ \varnothing \times \ \delta \ \delta \ w \ ln(1)os^{scB-5}y; P[sisC^+w^{+mC}]/+$. *Binsinscy* inversions duplicate *scB⁺*. Each cross included at least 100 *Dp[sisB⁺]/ln(1)os^{scB-5}; P[sisC⁺]/+* control females.

† $y \ w/w \ cm \ X^{f1} \ cf^B; P[sisA^+w^{+mC}]/A1 - B2 \ \& \ P[sisA^+w^{+mC}]/A1-H1/CyO \ \varnothing \ \varnothing \times \ \delta \ \delta \ (I)w/y; P[sisC^+w^{+mC}]/10-A/+; (II)w/y; P[sisC^+w^{+mC}]/10-B/+; (III)w/y; P[sisC^+w^{+mC}]/10-A \ \& \ P[sisC^+w^{+mC}]/10-B/+.$ Note that transgene 10-A = 89% and 10-B = 90% female rescue in part (q) above. Progeny were laid at 18°C and shifted to 25°C after at least one day.

‡ Viability of the reference male sib class itself measured relative to (50%) of their $Sxl^2/ISxl^2; P[sisA^+]/P[sisA^+]/+$ sisters was 99%, 95% and 86% for crosses I, II and III, respectively.

With the discovery of *sisC*, the collection of fly XSEs may be nearly complete^{4,5}. The impression given by this collection is that *Drosophila* relies on biochemically diverse proteins to assess X-chromosome dose, but they all act on *Sxl* at the level of transcription. In contrast, the XSEs of *Caenorhabditis elegans* include both transcriptional and post-transcriptional regulators of their target, *xol-1* (ref. 15). Characterization of *sisC* reveals that both *C. elegans* and *Drosophila* XSEs seem to include proteins that work extracellularly. □

hop⁻ and *Stat92E⁻* clones were generated by the FLP-mediated recombination dominant-female-sterile (FLP-DFS) technique¹⁶. The following crosses (markers omitted) generated females in which clones were induced and control populations of embryos: (1) *hop²* *P*[*FRT*014A – B]101/*Binsinsky*; *P*[*Sxl_{pe}*:*lacZ*]3.0 ♀♀ × ♂♂ *ova*^{D1} *P*[*FRT*]101/*Y*; *MKRS*, *P*[*hslFLP*]86*E* /*nkd*; and (2) *P*[*Sxl_{pe}* : *lacZ*]3.0; *Stat92E*^{6346(*P*:*lacZ*)} *P*[*FRT*]82*B*/TM3 ♀♀ × ♂♂ *P*[*hslFLP*]1/*Y*; *P*[*FRT*]82*B**P*[*ovo*^{D1}]18/TM3. Clone-bearing females were mated to wild-type males, so half their progeny carried *Sxl_{pe}:lacZ* (*n* = 218 and 437 for *hop* and *Stat92E* respectively, with *n* = 1,909 and 1,740 for controls).

Correspondence and requests for materials should be addressed to T.C. (e-mail: sxlcline@uclink.berkeley.edu). See GenBank accession no. AF231684 for *os* DNA.

Compromised HOXA5 function can limit p53 expression in human breast tumours

Venu Raman*, Shelby A. Martensen*, David Reisman†, Ella Evron*, Ward F. Odenwald‡, Elizabeth Jaffee*, Jeffrey Marks§ & Saraswati Sukumar*

* Breast Cancer Program, Johns Hopkins Oncology Center, Baltimore, Maryland 21231, USA

† University of South Carolina, Columbia, South Carolina 29208, USA

‡ National Institutes of Health, Bethesda, Maryland 20892, USA

§ Duke University Medical Center, Durham, North Carolina 27710, USA

Expression of the *p53* gene protects cells against malignant transformation^{1,2}. Whereas control of *p53* degradation has been a subject of intense scrutiny, little is known about the factors that regulate *p53* synthesis^{1,2}. Here we show that *p53* messenger RNA levels are low in a large proportion of breast tumours. Seeking potential regulators of *p53* transcription, we found consensus HOX binding sites^{3,4} in the *p53* promoter⁵. Transient transfection of Hox/HOXA5 activated the *p53* promoter. Expression of HOXA5 in epithelial cancer cells expressing wild-type *p53*, but not in isogenic variants lacking the *p53* gene⁶, led to apoptotic cell death. Moreover, breast cancer cell lines and patient tumours display a coordinate loss of *p53* and HOXA5 mRNA and protein expression. The HOXA5 promoter region was methylated in 16 out of 20 *p53*-negative breast tumour specimens. We conclude that loss of

expression of *p53* in human breast cancer may be primarily due to lack of expression of HOXA5.

Inactivation of *p53* caused by mutation is low (20%) in human breast cancer, as discussed in a recent meta-analysis⁷. Looking for other mechanisms that may account for loss of *p53* function in these tumours, we examined the levels of *p53* mRNA in breast cancer cell lines and in primary tumours. We found that *p53* mRNA levels were 5–10-fold lower in tumour cells than in normal breast epithelium (Fig. 1a, b). We then looked for consensus protein binding sites in the *p53* promoter⁵, including those of HOX proteins which are known to function as transcription factors^{8,9}. Selected HOX genes are differentially expressed in neoplasms of a number of tissues^{10–14}, but their functional relationship to the neoplastic phenotype remains to be elucidated. Six putative HOX-core binding sequences (ATTA) were identified within the 2.4-kilobase (kb) human *p53* promoter. Of a number of HOX genes examined in breast tumour cells and control breast epithelium, HOXA5 mRNA levels were markedly reduced in breast cancer cells (Fig. 1c). In fact, there was a tight correlation between *p53* and HOXA5 mRNA levels in the 10 cell lines tested for both genes, with correlation coefficient $r = 0.942$ (ref. 15). No such decreased expression was observed for HOXA10, B3, B7 or C8 mRNAs (data not shown).

To test for a causal relationship between the decreased expression of *p53* and HOXA5 mRNAs, we co-transfected ZR75.1 breast cancer cells or SAOS2 osteosarcoma cells with the –356 bp or the –2.4 kb human *p53* promoter–Luciferase reporter together with HOX expression plasmids. We observed that HOXA5 transactivated the *p53*-promoter-dependent reporter activity up to 25-fold in ZR75.1 cells and up to 7-fold in SAOS2 cells (Fig. 2a). This effect was not seen with the other homeotic genes *HOXB4*, *HOXB5* and *HOXB7* (Fig. 2a).

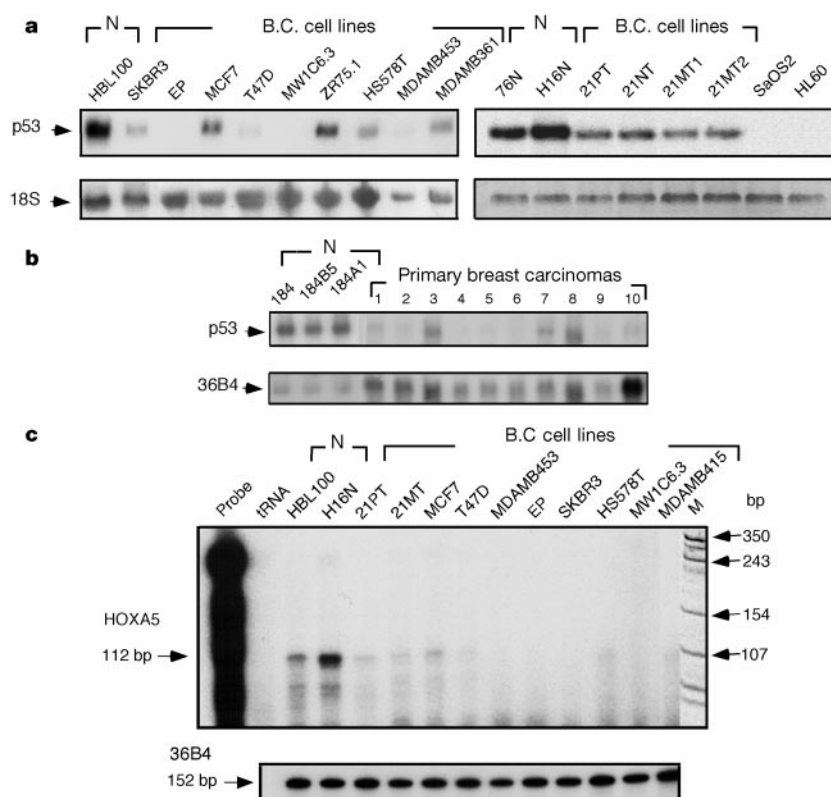


Figure 1 Human breast cancers express low levels of *p53* and HOXA5 mRNA. **a, b**, Northern blot analysis of normal (N) mammary epithelial cells: HBL100, 76N, H16N, 184, 184B5 and 184A1, and the indicated human breast cancer (B.C.) cell lines (**a**) and primary tumours (**b**). SAOS2 and HL60 cells, negative controls for *p53* mRNA expression.

c, Ribonuclease protection analysis of HOXA5 expression in normal epithelial (N) and breast cancer (B.C.) cell lines. Loading controls, 18S rRNA and 36B4 ribosomal protein gene.

We also observed positive regulation of transcription by HoxA5 with the mouse *p53* promoter. We found a single putative Hox-binding sequence (located at -204 to -201) in the upstream regulatory region of the murine *p53* gene. SAOS2 cells were cotransfected with a -320 bp mouse *p53* promoter fused to the CAT gene, together with expression plasmids encoding full-length murine HoxA5, HoxA7 or HoxC8 proteins. Similar to human HOXA5, we found a 15–20-fold increase in CAT activity in the SAOS2 cells co-transfected with the HoxA5 expression plasmid, but no significant effect with HoxA7 or HoxC8 (Fig. 2b). These results indicate that expression from the mouse *p53* promoter may be specifically stimulated by HoxA5. To define the sequence requirements for the transactivation function, we tested a deletion construct of the *p53*-promoter CAT construct (Fig. 2c) in co-transfection assays with the full-length HoxA5 expression plasmid. A deletion to -153 bp in the promoter region of the *p53*-CAT construct eliminated stimulation of CAT activity by the effector plasmid (Fig. 2d). A truncated HoxA5 protein termed pCMVΔHoxA5, which lacks the homeodomain, was completely inactive in these experiments (Fig. 2d). Finally a 'TT' to 'GG' mutation in the core binding site (-320 mp53Mut CAT) that abolished DNA–protein complex formation in cell extracts (see

below), completely abrogated transactivation of the CAT reporter gene by HoxA5 (Fig. 2d).

Direct binding of HoxA5 to the ATTA-containing site in the *p53* promoter (positions -204 to -201) was demonstrated by electrophoretic mobility shift (EMSA) and supershift assays. A band was observed in cell extracts from cells transfected with HoxA5, but not in extracts from control cells (Fig. 2e, lane 2). This band was competed out by an excess of unlabelled oligonucleotide but not by an oligonucleotide with an unrelated sequence (Fig. 2e). No protein–DNA complex was observed in extracts mixed with an oligonucleotide primer that carries two mutations (TT to GG) in the core binding site. Finally, HOXA5 antibodies (Fig. 2f, lane 4), but not pre-immune serum, caused a supershift of the bound HOXA5 protein–oligonucleotide complex (Fig. 2f). This supershift was abrogated by pre-incubation with excess antibody (antigen depletion) (Fig. 2f). Similar shift patterns were observed in extracts of RKO cells transfected with the effector plasmid (data not shown). These results indicate that the ATTA-containing sequence in the mouse *p53* promoter is indeed a HoxA5-binding sequence.

The above results suggest that HOXA5 may possess growth-suppressive properties through activation of *p53* expression. To test this possibility, breast cancer cells MCF-7 and ZR75.1, which

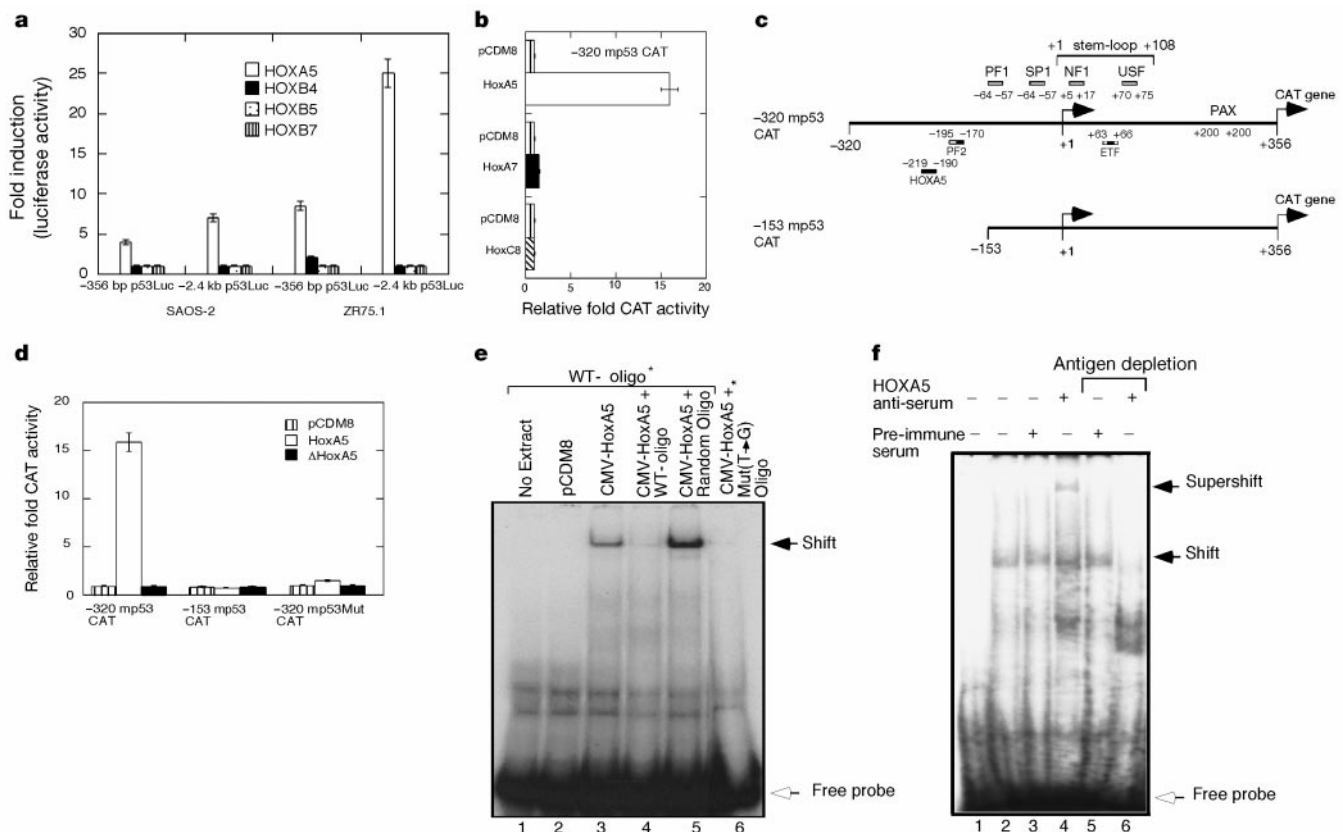


Figure 2 Mouse and human HOXA5 transactivate the *p53* promoter. **a**, Luciferase assays on extracts of SAOS2 or ZR75.1 cells transfected with human HOXA5, HOXB4, HOXB5 or HOXB7 expression plasmids plus human -2.4 kb or -356 bp *p53* promoter–Luc constructs. The relative fold increase in activity compared with cells transfected with vector or Renilla luciferase plasmid was determined. Mean $\pm$ s.d. of four independent experiments is shown. **b**, Mouse -320 mp53 CAT reporter plasmid was co-transfected into SAOS2 cells with the mouse HoxA5, HoxA7 or HoxC8 expression vector or empty vector, pCDM8. Mean $\pm$ s.d. of three independent assays²⁷ is shown. **c**, Representation of the -320 mp53 CAT and its deletion construct, -153 mp53 CAT. **d**, SAOS2 cells were transiently transfected with the constructs -320 mp53 CAT, -153 mp53 CAT or -320 mp53 CAT with a TT to GG mutation in the core Hox-binding site (-320 mp53Mut CAT), along with the expression vector pCDM8, or with pCMVHoxA5 or pCMVΔHoxA5. Mean $\pm$ s.d. of five experiments is shown. **e**, For EMSA²⁸, end-labelled wild type (WT)

oligonucleotide probe incubated with no extract (lane 1) or extracts (3 μ g protein) from SAOS2 cells transfected with vector alone (lane 2) or pCMVHoxA5 alone shows a distinct band (lane 3), which is competed out by a 50-fold excess of unlabelled wild-type oligonucleotide (lane 4) but not by an oligonucleotide of unrelated sequence (lane 5). Lane 6, mutant oligo does not bind HoxA5 in the protein extract from cells transfected with pCMVHoxA5. Asterisk, double-stranded wild-type (5'-AATGCTATTTT GAATTAAGAAAGGTGAGA-3') or mutant (5'-AATGCTATTTTGAAGGAAGAAAGGTGAGA-3') oligonucleotides end-labelled with [γ -³²P]ATP. **f**, Supershift of the bound HoxA5 protein–oligonucleotide complex by HOXA5 antibodies (lane 4), but not the pre-immune serum (lane 3), abrogated by pre-incubation of the cell extract with excess antibody (antigen depletion) (lanes 5, 6). Lane 1, no extract; lane 2, cell extract plus oligonucleotide without anti-serum.

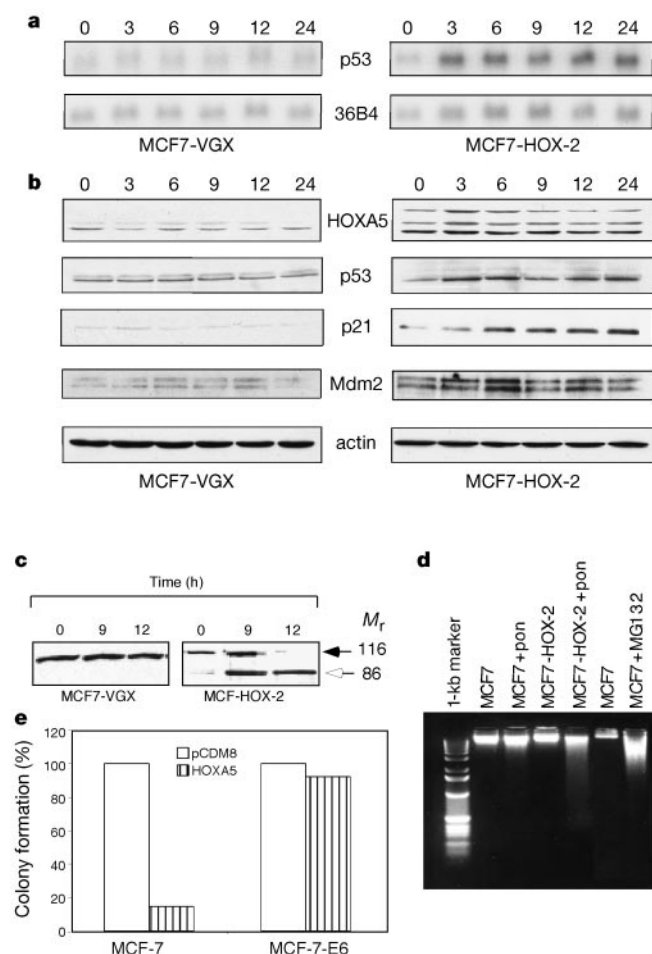


Figure 3 HOXA5 upregulates p53 levels and initiates an apoptotic death pathway in breast cancer cells. **a**, Northern analysis shows induction of p53 mRNA expression in clone MCF7-HOX-2 after addition of Ponasterone A (5 μ M), but not in cells expressing the vector pVgRXR. **b**, Western blot analysis of cell extracts at different times after addition of Pon A to cultures of MCF7-HOX-2. HOXA5 is frequently seen as a relative molecular mass (M_r) 40–42 doublet; a third band is seen after Pon A induction of the transfected HOXA5 gene in MCF7-HOX2 cells. The multiple bands may be due to post-translational modifications or alternate splicing. **c**, DNA fragmentation analysis. DNA stained by ethidium bromide, after treatment for 48 h with MG132 or Pon A. **d**, Western blot for PARP cleavage¹⁶. Shown are extracts of MCF7-VGX and MCF7-HOX-2 cells after exposure to Pon A for the indicated times in culture. **e**, Colony formation assay of MCF-7 and MCF-7-E6, 2 weeks after selection of pCMVHOXA5 transfected cells for blasticidin resistance.

have wild-type *p53* genes, were transfected with the full-length HOXA5 and the Δ HOXA5 (homeodomain-deleted) expression plasmids and tested for their colony-forming ability. No surviving colonies were obtained from HOXA5-transfected cells whereas those transfected with Δ HOXA5 and the vector control generated colonies with equal efficiency (data not shown). To obtain stable cultures that could express HOXA5, we generated clones of MCF-7 cells in which the HOXA5 gene was placed under the control of an ecdysone-inducible promoter. Less than 3 h after induction of HOXA5 expression by the ecdysone analogue, Ponasterone A (Pon A), the levels of p53 mRNA rose twofold (Fig. 3a). Western blotting showed that p53 and its downstream targets, p21 and Mdm2, as well as HOXA5 were reproducibly induced 2–5-fold after treatment with Pon A (Fig. 3b). Moreover, by 24 h Pon A had caused cell shrinkage, which was followed by significant cell death (80–90%) after 48 h. Cell death occurred by apoptosis according to the following criteria¹⁶: first, cells shrank and formed contractile

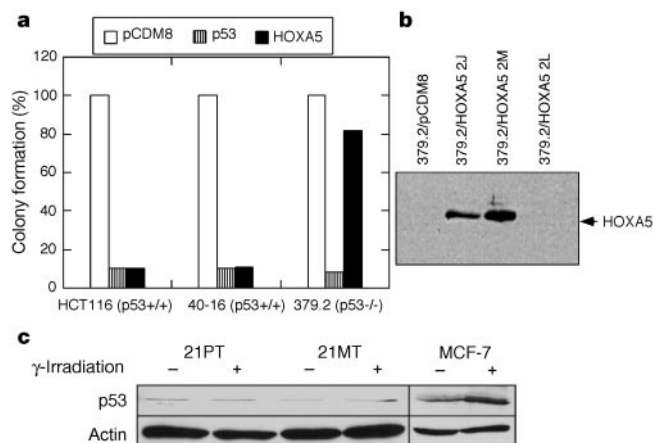


Figure 4 HOXA5 induces cell death in *p53*^{+/+} cells but not in *p53*^{-/-} cells. **a**, Colony formation assays of HCT116 colon carcinoma cells, and clones 40-16 (*p53*^{+/+}) and 379.2 (*p53*^{-/-}) cells, co-transfected with the vector (pCDM8), p53 (pCMVp53) or with HOXA5 (pCMVHOXA5), along with the blasticidin selection plasmid (pBOSH2BGFP). The number of colonies varied between 290 and 400 colonies per plate in the vector-transfected HCT-116 parental cells and between 30 and 60 colonies per plate in the clones 40-16 and 379.2. The assay was repeated five times, with $\leq 10\%$ variation in colony number between experiments. **b**, Western analysis of clones picked from 379.2 cells transfected with pCMVHOXA5 show the expression of the M_r 40–42 HOXA5 protein in 2 out of 3 transfectants. **c**, γ -Irradiation does not induce p53 in 21PT or 21MT cells. One million cells were seeded in 10-cm tissue-culture dishes. After 24 h, cells were treated with 10 Gy of γ -radiation. Cells were harvested after 6 h and analysed for p53 expression by western blotting as described in Methods.

bodies (data not shown); second, DNA laddering was observed (Fig. 3c); third, poly(ADP-ribose) polymerase, a substrate for caspases, underwent cleavage by 12 h (Fig. 3d); and last, 70% of the cells showed micronucleus formation, membrane blebbing and ghost cell features upon staining with acridine orange (data not shown). This apoptosis was not accompanied by a detectable change in the levels of Bax protein (data not shown).

The above results are consistent with the hypothesis that an increase in the level of HOXA5 in MCF-7 cells leads to an increase in p53 levels, which in turn results in apoptosis. As a further proof of this model, MCF-7 cells expressing the E6 gene of human papilloma virus, when transfected with the HOXA5 expression vector, were fully able to form colonies (Fig. 3e). Presumably, the induced p53 in these cells was sequestered by E6 protein and was unable to induce apoptosis. These results support the idea that HOXA5 induces apoptosis through a p53-dependent pathway in MCF-7 cells. To our knowledge, this is the first demonstration of the involvement of a HOX protein in apoptosis.

We also tested the idea that HOXA5-induced apoptosis is mediated by p53. We transfected *p53*^{+/+} HCT 116 line of colon carcinoma cells and its *p53*^{-/-} derivative clone 379.2 (ref. 6) with HOXA5 and p53 expression vectors. Expression of HOXA5 or p53 in the parental HCT116 cells reduced the ability of the cells to form colonies (Fig. 4a). In contrast, HOXA5 and p53 expression led to different phenotypes in *p53* null 379.2 cells. Whereas expression of p53 in these cells abrogated colony formation, expression of HOXA5 had no detectable effect. In the HOXA5-transfected cultures, stable colonies, expressing detectable amounts of HOXA5 protein (Fig. 4b), and of a size and number comparable to the vector control were observed. Thus, HOXA5 induces cell death only in the presence of a wild-type *p53* gene, adding further evidence that p53 mediates HOXA5 activity. Conversely, cells lacking HOXA5 and p53 would be unable to mount a normal response to treatments, such as DNA damage, that normally raise p53 levels by stabilizing the

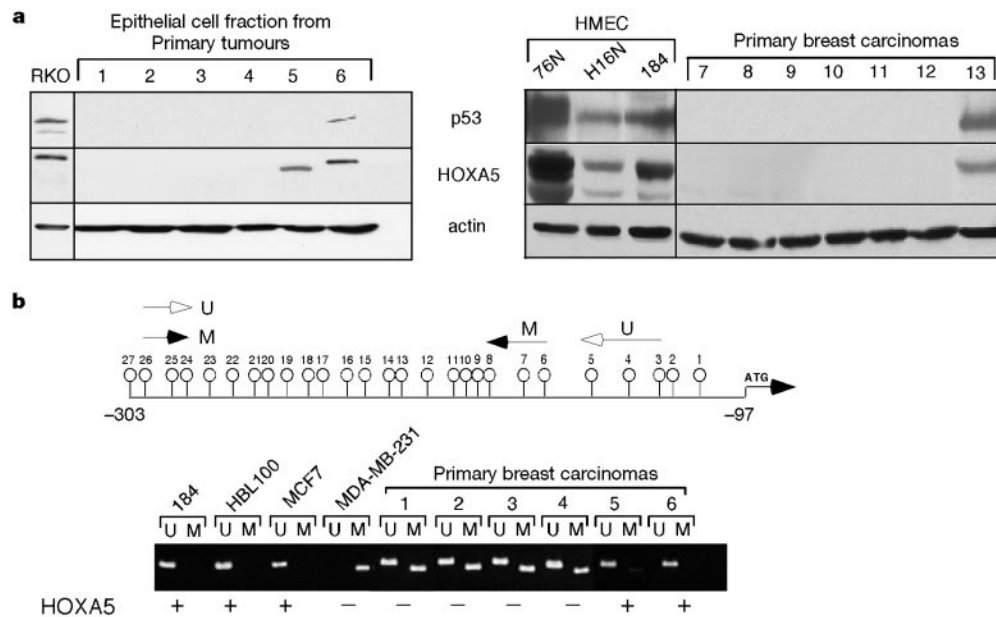


Figure 5 Loss of HOXA5 expression in breast cancer—gene silencing by methylation. **a**, Concurrent loss of HOXA5 and p53 protein expression in primary breast carcinomas by western analysis using the HOXA5 and p53 antibodies described above. Colon carcinoma cells (RKO), containing wild-type p53 and HOXA5, were used as controls. Normal mammary epithelial cells (HMEC) and 13 representative primary breast cancers out of 30 that were tested. **b**, Methylation-specific PCR (MSP) analysis of sodium-bisulphite-treated

DNA from cultured finite lifespan HMEC-184, an immortal HMEC-HBL100, breast cancer cell lines MCF-7 and MDA-MB-231, and primary carcinomas 1–6. Representation of the HOXA5 promoter region containing 27 CpGs, with the location of the primers used for amplifying methylated (M) and unmethylated (U) DNA is shown above. The HOXA5 expression data of the tumours from the western blots (**a**) is summarized below.

protein. To test this possibility, we treated the two tumour cell lines 21PT and 21MT, which have low expression of HOXA5 and p53 (see Fig. 1a, c), with γ -radiation. No detectable increase in p53 level in 21 PT and 21 MT was observed, whereas, as expected, p53 was induced in MCF-7 cells (Fig. 4c).

Our findings from cell-culture experiments had *in vivo* correlates. In 67% (20 out of 30) of primary breast tumours, HOXA5 protein was undetectable (Fig. 5a). Strikingly, concurrent loss of p53 expression was observed in the same tumours that lacked HOXA5 (Fig. 5a). Among those tumours expressing HOXA5, one (lane 5) showed a band migrating faster than the wild-type HOXA5 in the RKO cells (Fig. 5a). We sequenced HOXA5 cDNAs from 11 p53-negative breast cancer samples and 2 finite lifespan human breast epithelial cell (HMEC) strains. All HOXA5 coding regions were wild type except that of tumour 5, which contained a frameshift mutation (G insertion at codon 204) that created a premature stop codon. There is a coupled loss of p53 and HOXA5 expression in primary breast carcinomas, possibly due to lack of expression or mutational inactivation of HOXA5.

Seeking a reason for the absence of the protein in the tumours, we sequenced HOXA5 DNA of 20 HOXA5-negative and 5 HOXA5-positive primary tumours. All contained the wild-type sequence except tumour 5, in which the insertion of G was again found and which contained no wild-type allele (data not shown). In the absence of mutations, loss of HOXA5 may be a consequence of a loss of upstream regulatory factors^{17–19} or may reflect some repressive phenomenon such as methylation of the gene²⁰. Methylation-specific polymerase chain reaction (MSP) of sodium-bisulphite-treated DNA²¹ showed that 16 out of 20 of the tumours contained partially or completely methylated CpGs in the HOXA5 promoter region (accession number AC004080) (Fig. 5b). In contrast, this region was completely unmethylated in human mammary epithelial cells (HMEC) of finite lifespan, 184 and 9F1403, and in 4 immortalized HMECs, HBL100, MCF10A, 184B5 and 184A1 (representative sample shown in Fig. 5b). Nucleotide sequencing of the region –97 bp to –303 bp of the HOXA5 promoter (shown in

Fig. 5b), using sodium bisulphite-treated DNA from HMEC 184, and cancer cell lines MCF-7 and MDA-MB-231, showed that methylation correlated with silencing of gene expression (data not shown). Expression of HOXA5 mRNA could be re-initiated in MDA-MB-231 cells by treatment with the DNA methyl transferase inhibitor, 5-aza-2'-deoxycytidine (5-aza-dC) (data not shown). These results are strong preliminary evidence that methylation of the HOXA5 promoter region may be responsible for silencing of gene expression.

What is the biological significance of the positive transcriptional regulation of p53 by HOXA5? Unlike most tumour types^{1,2}, up to 80% of sporadic breast cancers do not contain p53 mutations^{7,22,23}. Our results suggest that the reduced p53 levels in these tumours result from the absence of a positive regulator of p53 mRNA synthesis^{1,2,24}. p53 normally functions as a tetramer, so even a small reduction in the concentration of p53 monomers can greatly reduce the effective concentration of tetramers^{25,26}. We have shown that transfected HOXA5 upregulates both p53 promoter-reporter constructs and endogenous p53 synthesis, leading to apoptosis. Finally, HOXA5 was detectable in only one-third of the primary tumours. In most of the remaining tumours, lack of HOXA5 expression strongly correlated with methylation of its promoter region, suggesting a causal role for methylation in the silencing of HOXA5 gene expression.

In summary, our experiments show that HOXA5 is a positive regulator of p53 transcription and function in cultured cells. The correlations we have observed between HOXA5 and p53 levels in clinical breast cancer strongly suggest that loss of HOXA5 expression is an important step in tumorigenesis. □

Methods

Plasmid constructions and reporter assays

The human p53–luciferase constructs were generated by excising the 2.4-kb *XbaI* fragment and the 356-bp *XbaI*–*Bam*HI fragments of p53 from pICAT (ref. 2), and subcloning each into the *SmaI* site of pGL2 basic luciferase reporter vector from Promega. For the construction of the mouse, –320 mp53 CAT, a 676-bp *Hind*III–*Eco*RI restriction fragment

from a mouse genomic clone was subcloned into an engineered *SaI* cloning site in the vector pUC18CAT. The -153 mp53 CAT clone was obtained from -320 mp53 CAT by *Bal31* digestion. For transfection assays, the indicated cancer cells were seeded (5×10^5 cells) in 60-mm tissue-culture dishes in DMEM supplemented with 10% fetal bovine serum 24 h before transfection. The effector and reporter plasmids were transfected using lipofectamine according to the manufacturer's instructions (Panvera). Luciferase and CAT activities were assayed 36 h after transfection. For luciferase assays, the relative activity compared with vector-transfected cells and Renilla luciferase plasmid was determined for each experiment. For CAT assays, transcriptional activity was expressed as conversion of chloramphenicol to acetylated chloramphenicol relative to that in cells transfected with the -320 mp53 CAT alone without the expression vector, normalized for protein concentration of the extracts²⁷.

Tissue preparations and cells

Freshly excised primary breast carcinomas or mammaplasty specimens were minced finely with razor blades and digested with 0.15% collagenase A and 0.5% dispase II (Boehringer Mannheim) prepared in RPMI 1640 medium. The cell clumps were separated from the lighter fibroblasts by gravity separation three times. The cell clumps were then digested for 15 min with trypsin, washed and immunostained with anti-cytokeratin-specific antibody (CAM 5.2, Becton-Dickinson) to assess the level of epithelial cell enrichment. The epithelial cells comprised between 70 and 80% of the enriched cell population.

Frozen, surgically excised breast tumour samples were cryosectioned, and representative sections were screened by a pathologist after staining with haematoxylin and eosin. Sections containing more than 70% carcinoma cells were used for RNA and protein extractions directly. Breast cancer cell lines and immortalized HMECs were obtained from ATCC (Rockville, MD). Finite lifespan HMECs were obtained from M. Stampfer, HMEC strain 9F1403 from Clonetics.

Methylation-specific PCR (MSP) and sodium bisulphite DNA sequencing

Genomic DNA (1 µg) was treated with sodium bisulphite²¹ and was analysed for MSP using primer sets specific for methylated DNA, 5'-TTTAGCGGTGGCGTTCG-3' (sense) and 5'-ATACGACTTCGAATCACGTA-3' (antisense); and primers specific for unmethylated DNA, 5'-TTGGTGAAGTTGGGTG-3' (sense), and 5'-AATACAACCTTCAAATCAC ATAC-3', which yielded products of 183 and 213 bp, respectively. Sodium-bisulphite-treated DNA was used to PCR-amplify the HOXA5 promoter region -97 to -303 bp, using the primers 5'-ATTTTGTATAATGGGTGTAAT-3' (sense) and 5'-AACATATACTTAA TTCCCTCC-3' (antisense). The product was purified using a Qiagen PCR purification kit (Qiagen) and was sequenced²¹ using the sense primer with an ABI automated fluorescent sequencer according to the manufacturer's instructions.

Treatment of cells with 5'-aza-2'-deoxycytidine (5-aza-dC)

MDA-MB-231 breast cancer cells were treated with 0.75 µM 5-aza-dC (Sigma)²¹, and collected 0, 3 and 5 days later. PCR with reverse transcriptase²¹ was performed using primers 5'-TCATTTTG CGGTCGCTATCC-3' (sense) and 5'-GCCGGCTGGCTGTACCTG-3' (antisense).

Immunoblot analysis

Proteins were visualized by western analysis and 10% SDS-polyacrylamide gel electrophoresis. The primary antibodies (anti-HOXA5 (HOXA5-2, BABCO), anti-p53 (AB-6, Oncogene Science), or anti-β-actin (AC-15, Sigma), anti-p21 (15091A, Pharmingen), anti-Mdm2 (65101A, Pharmingen), anti-PARP (AB-2, Oncogene Sciences), anti-dynein (Zymed) and anti-Na⁺,K⁺-ATPase (Ed Benz, Johns Hopkins) (also used as loading controls, with actin)) were used at 1:1000 dilution.

Received 13 March; accepted 9 May 2000.

1. Prives, C. & Hall, P. A. The p53 pathway. *J. Pathol.* **187**, 112–126 (1999).
2. Reisman, D. & Loging, W. T. Transcriptional regulation of the p53 tumor suppressor gene. *Semin. Cancer Biol.* **8**, 317–324 (1998).
3. Pellerin, L., Schnabel, C., Catron, K. M. & Abate, C. Hox proteins have different affinities for a

- consensus DNA site that correlate with the positions of their genes on the hox cluster. *Mol. Cell. Biol.* **14**, 4532–4545 (1994).
4. Zhao, J. J., Lazzarini, R. A. & Pick, L. Functional dissection of the mouse Hox-a5 gene. *EMBO J.* **15**, 1313–1322 (1996).
5. Reisman, D., Greenberg, M. & Rotter, V. Human p53 oncogene contains one promoter upstream of exon 1 and a second, stronger promoter within intron 1. *Proc. Natl Acad. Sci. USA* **85**, 5146–5150 (1988).
6. Bunz, F. *et al.* Requirement for p53 and p21 to sustain G2 arrest after DNA damage. *Science* **282**, 1497–1501 (1998).
7. Pharoah, P. D. P., Day, N. E. & Caldas, C. Somatic mutations in the p53 gene and prognosis in breast cancer: a meta-analysis. *Brit. J. Cancer* **80**, 1968–1973 (1999).
8. Deschamps, J. & Meijlink, F. Mammalian homeobox genes in normal development and neoplasia. *Crit. Rev. Oncol.* **3**, 117–173 (1992).
9. Scott, M. P. Hox genes, arms and the man. *Nature Genet.* **15**, 117–118 (1997).
10. Cillo, C. *et al.* Hox gene expression in normal and neoplastic human kidney. *Int. J. Cancer* **51**, 892–897 (1992).
11. De Vita, G. *et al.* Expression of homeobox-containing genes in primary and metastatic colorectal cancer. *Eur. J. Cancer* **6**, 887–893 (1993).
12. Care, A. *et al.* HOXB7 constitutively activates fibroblast growth factor in melanomas. *Mol. Cell. Biol.* **16**, 4842–4851 (1996).
13. Chariot, A., & Castronovo, V. Detection of HOXA1 expression in human breast cancer. *Biochem. Biophys. Res. Commun.* **222**, 292–297 (1996).
14. Cillo, C., Faiella, A., Cantile, M. & Boncinelli, E. Homeobox genes and cancer. *Exp. Cell. Res.* **248**, 1–9 (1999).
15. Snedecor G. W. & Cochran, W. G. *Statistical Methods* (Iowa State Univ. Press, Ames, 1982).
16. Li, F. *et al.* Cell-specific induction of apoptosis by microinjection of cytochrome c. Bcl-xL has activity independent of cytochrome c release. *J. Biol. Chem.* **272**, 30299–30305 (1997).
17. Laroche, C., Tremblay, M., Bernier, D., Aubin, J. & Jeanotte, L. Multiple cis-acting regions are required for restricted spatio-temporal Hoxa5 gene expression. *Dev. Dyn.* **214**, 127–140 (1999).
18. Aubin, J., Lemieux, M., Tremblay, M., Behringer, R. R. & Jeanotte, L. Transcriptional interferences at the Hoxa4/Hoxa5 locus: importance of correct Hoxa5 expression for proper specification of the axial skeletal. *Dev. Dyn.* **212**, 141–156 (1999).
19. Hanson, R. D. *et al.* Mammalian Trithorax and polycomb-group homologues are antagonistic regulators of homeotic development. *Proc. Natl Acad. Sci. USA* **96**, 14372–14377 (1999).
20. Bird, A. P. & Wolffe, A. P. Methylation-induced repression—belts, braces, and chromatin. *Cell* **99**, 451–454 (1999).
21. Herman, J. G., Graff, J. R., Myohanen, S., Nelkin, B. D. & Baylin, S. B. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc. Natl Acad. Sci. USA* **93**, 9821–9826 (1996).
22. Runnebaum, I. B., Nagarajan, M., Bowman, M., Soto, D. & Sukumar, S. Mutations in p53 as potential molecular markers for human breast cancer. *Proc. Natl Acad. Sci. USA* **88**, 10657–10661 (1991).
23. Glebov, O. K., McKenzie, K. E., White, C. A. & Sukumar, S. Frequent p53 gene mutations and novel alleles in familial breast cancer. *Cancer Res.* **54**, 3703–3709 (1994).
24. Bates, S. A. V. & Vousden, K. H. Mechanisms of p53-mediated apoptosis. *Cell. Mol. Life Sci.* **55**, 28–37 (1999).
25. Venkatachalam, S. *et al.* Retention of wild-type p53 in tumors from p53 heterozygous mice: reduction of p53 dosage can promote cancer formation. *EMBO J.* **17**, 4657–4667 (1998).
26. Gottlieb, E. *et al.* Transgenic mouse model for studying the transcriptional activity of the p53 protein: age and tissue dependent changes in radiation-induced activation during embryogenesis. *EMBO J.* **16**, 1381–1390 (1997).
27. Seed B. & Sheen, J. Y. A simple phase extraction assay for chloramphenicol acetyltransferase activity. *Gene* **67**, 271–277 (1988).
28. Raman, V., Andrews, M. E., Harkey, M. A. & Raff, R. A. Protein–DNA interactions at putative regulatory regions of two coordinately expressed genes, msp130 and PM27, during skeletogenesis in sea urchin embryos. *Int. J. Dev. Biol.* **37**, 499–507 (1993).

Acknowledgements

This work was supported by grants from the USPHS, the Susan G. Komen Foundation, the Department of Defense and the Breast Cancer Research Foundation to S.S. We thank A. Rein and B. Vogelstein for discussions and for critically reviewing the manuscript, V. Band for generously sharing data and reagents, and D. Korz for expert technical assistance.

Correspondence and requests for materials should be addressed to S.S. (e-mail: saras@jhmi.edu).

Instrumental in analysis

Analytical instruments, from electric nose to GasHound.

MA50/100 analysers

Sartorius www.sartorius.com

Flexible moisture analysis

In this pair of new moisture analysers, the heating unit can be changed to suit a particular application. The ceramic radiator is for ultra-fast and uniform heating and to avoid overheating of specific areas of a sample. The halogen lamp unit is for flexibility of ease of use. The motorized heating unit is designed for easy access to the sample chamber, whilst minimizing the vibration caused by opening and closing. The 0.1 mg resolution of the MA100, with a weighing capacity of 100 g, is equivalent to the accuracy of an analytical balance.

Reader Service No. 100



Weigh upstream: the Sartorius moisture analyser.

FTIRsearch.com

Galactic Industries/Nicolet

www.ftirsearch.com

Pay-per-use searching of FTIR and Raman spectral libraries

This online service allows analytical chemists to search FTIR and Raman spectral libraries via the Internet. Launched on 17 April this year, FTIRsearch.com offers instant access to over 71,000 FTIR and almost 16,000 Raman spectra. Users can perform both Text and Spectral searching of the complete database.

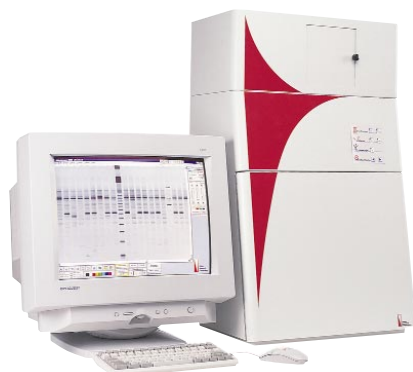
Reader Service No. 101

FluorChem 8000

Alpha Innotech www.alphainnotech.com

High-density array imaging

The FluorChem 8000 Advanced Fluorescence, Chemiluminescence and Visible Imaging System is recommended for high-density array fluorescence, chemiluminescence and colorimetric blots. Resolution is over 1.3 million pixels, quantum efficiency is nearly 60% at 400 nm. The system is also



Alpha particles: FluorChem 8000.

available combined with the AlphaEase FC software package of analysis tools.

Reader Service No. 102

Emission Analyser

Nordic Sensors Technology

www.nordicsensor.se

Detect odours and gas emissions

The electronic nose from Nordic Sensors Technology has been revised for higher sensitivity, speed and accuracy. The analyser consists of an array of electronic chemical sensors that exhibit partial specificity to a variety of gases and input their responses to a computerized pattern recognition system. In the latest model the sensors are repositioned closer to the point of entry of the sample emission, the software is revised and there is now the option of a stand-alone or online instrument.

Reader Service No. 103

Amalgam-2

Hitachi

www.hsi-europe.com

Mercurial analysis

Hitachi Scientific Instruments has introduced a new system to reduce the difficulties associated with detecting mercury at very low concentrations by atomic absorption spectrophotometry. The Amalgam-2 system is used in conjunction with the Z-5000 AA spectrophotometer and the HFS-3 hydride system. A detection limit for mercury of the order of parts per trillion can be achieved. Detection levels for elements such as Hg, As, Sb, Se, Sn and Bi have traditionally been reduced for flame AAS using the hydride generation method. Amalgam-2 is a self-contained concentration unit which is located between the HFS-3 hydride generator and the Z-5000 AA spectrophotometer. Hydrogen is produced in the HFS-3 by the reaction of sodium borohydride with hydrochloric acid and this reduces

mercury-containing samples to form elemental mercury. The Amalgam-2 system is placed in the gas stream carrying the cold mercury vapour from the HFS-3 to the Z-5000. It consists of a concentration cell which is a capillary filled with gold wool. This collects the mercury as amalgam.

Reader Service No. 104

Fluorescence standards

Matech

www.pdcorp.com

A measured response

The Matech Fluorescent Reference Standard distributed by Precision Dynamics Corporation (PDC) allows you to calibrate wavelengths and intensity according to your own requirements. The makers claim that for both wavelengths and intensity calibration these inorganic fluorescent reference standards offer stability superior to organic dye doped polymer standards or quinine sulphate (ASTM method). The standards come in 96-well microtitre plate format for simple calibration of assay instrumentation such as fluorescent plate readers and DNA sequence detectors.

Reader Service No. 105

Gashound

Li-Cor

www.licor.com

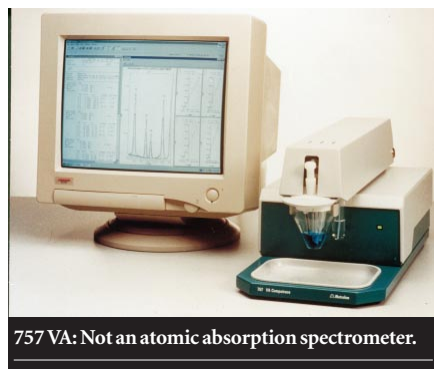
Precision CO₂ analysis

GasHound is a single-path, dual-wavelength, non-dispersive infrared gas analyser designed for quiet operation. Features include automatic temperature and pressure compensation, and interchangeable optical benches. The machine operates on two ranges, 0–2,000 p.p.m. or 0–20,000 p.p.m. The unit is small and light in weight and can be operated from mains, batteries or most DC power supplies.

Reader Service No. 106



Gas guzzler: Li-Cor's GasHound on the scent.



757 VA: Not an atomic absorption spectrometer.

757 VA Computrace

Metrohm www.metrohm.co.uk

A trace analyser with determination

This PC-controlled analyser operates under Windows with a choice of two modes. The 'Determination' mode is for routine analysis of sub-parts per billion in metals. And 'Exploratory' is for method development and research. The 757 can connect up to 5×665 Dosimats, making buffer and standard additions automatic, and a library of standard application bulletins. Metrohm claim that the 757 VA Computrace does the work of an atomic absorption spectrometer, at half the cost.

Reader Service No. 107

Portable Analytical System

MKS Instruments www.mksinst.com

Monitoring system for process chambers

The Portable Analytical System (PAS) from MKS Instruments now has the capability to monitor process chambers at process pressures up to 10 torr. It incorporates the new ORION process monitor Residual Gas Analyser and the system can now continuously acquire data for up to 10 days. The ORION software monitors the levels of residual contaminants in the process chamber and automatically alerts the operator if process parameters are out of specification. The software can also be set up to continually monitor for air leaks, internal water leaks and other occurrences in the process chamber whilst simultaneously operating as a helium leak detector.

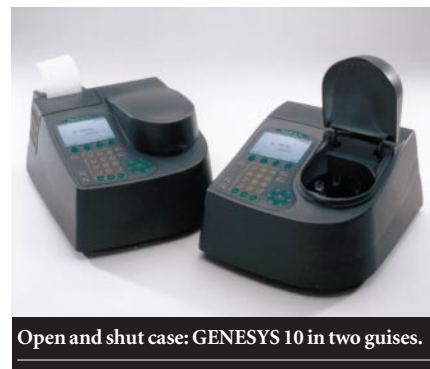
Reader Service No. 108

GENESYS 10

Spectronic Unicam www.spectronic.com

UV-visible spectrophotometers

Available in both visible (325–1,100 nm) and UV-visible (190–1,100 nm) models, the GENESYS 10 spectrophotometer features a 5-nm spectral bandwidth. The imaging sys-



Open and shut case: GENESYS 10 in two guises.

tem allows the use of microcells without the need for special cell holders. The UV-visible model features a xenon lamp, dual detectors and split-beam optics to maximize signal-to-noise ratio. The machine comes pre-programmed to run complex calculations and contains eight built-in software programs including Standard Curve, Absorbance Ratio, Survey Scan, Kinetics and Performance Validation for fast, efficient data processing.

Reader Service No. 109

These notes are compiled in the Nature office from information provided by the manufacturers. For more details, fill in the reader service card.

ADVERTISEMENTS

**BACs • YACs • PACs
cDNAs**

Research Genetics provides clones and colony membranes for a number of human, mouse and rat libraries as well as a custom screening service for the libraries.

Research Genetics
U.S. or Canada 800-533-4363
U.K. 0-800-89-1393
FAX 256-536-9016

Check out our homepage at
<http://www.resgen.com>

READER ENQUIRY NO. 40

Look no further
for the best value in
Oligo Synthesis

Competitive pricing
Routine 24 hour dispatch • Every Oligo QC verified by PAGE
Rigorous batch QC by PCR and Mass Spec
Web ordering and tracking • Local Freephone numbers
A great deal in any language!!

SIGMA (+44) (0) 1223 839000 • fax (+44) (0) 1223 839200
GENOSYS email: info@sigma-genosys.co.uk • www.sigma-genosys.co.uk

READER ENQUIRY NO. 33

World's First Universal Quantum Computation Simulator
**QUANTUM COMPUTER
SIMULATOR**
for Windows/Macintosh/UNIX

The DEMO version is freely downloadable at our Web site. Many sample algorithms including Shor's factorization algorithm and Grover's database search algorithm which are most popular in the field of quantum computer study and manuals are also available with free of charge.

<http://www.senko-corp.co.jp/qcs>

SENKO
CORPORATION

9-30, Kitasaiwai 2-chome, Nishi-ku,
Yokohama, 220-0004 Japan
PHONE: 81-45-320-2711
FAX: 81-45-320-9342
e-mail: senko@246.ne.jp

READER ENQUIRY NO. 44

Spektakulär 089

was wissen schaf f

Das grosse Fachbuch-Programm bei www.amazon.de.
Mit der aktuellen Literatur für Bio-, Geo- und Ingenieurwissen-
schaften, Medizin, Chemie, Mathematik, Physik oder Astronomie.
Fachspezifisch strukturiert und ganz schnell gefunden.

www.amazon.de

einfach so einkaufen
amazon.

READER ENQUIRY NO. 46